
 Europäisches Patentamt  
 European Patent Office  
 Office européen des brevets

(11) Publication number:

0 251 446  
A2

(12)

EUROPEAN PATENT APPLICATION

(21) Application number: 87303761.8

(51) Int. Cl.<sup>3</sup>: C 12 N 15/00  
C 12 N 9/54, C 12 N 1/00

(22) Date of filing: 28.04.87

(30) Priority: 30.04.86 US 858594  
06.04.87 US 35652

(43) Date of publication of application:  
07.01.88 Bulletin 88/1

(84) Designated Contracting States:  
AT BE CH DE ES FR GB GR IT LI LU NL SE

(71) Applicant: GENENTECH, INC.  
460 Point San Bruno Boulevard  
South San Francisco California 94080(US)

(72) Inventor: Wells, James Allen  
64 Otay Avenue  
San Mateo CA 94403(US)

(72) Inventor: Cunningham, Brian C.  
24 Olive Avenue  
Piedmont CA 94611(US)

(72) Inventor: Caldwell, Robert Mark  
1828 Broadway  
No.101 San Francisco Ca 94109(US)

(72) Inventor: Bott, Richard Ray  
3032 Hillside drive  
Burlingame CA 94010(US)

(72) Inventor: Estell, David Aaron  
250 Diablo Avenue  
Mountain View CA 94043(US)

(72) Inventor: Power, Scott Douglas  
732 Olive Court  
San Bruno CA 94066(US)

(74) Representative: Bizley, Richard Edward et al,  
BOULT, WADE & TENNANT 27 Fenchurch Street  
London EC4A 1PQ(GB)

EP 0 251 446 A2

(54) Non-human Carbonyl hydrolase mutants, DNA sequences and vectors encoding same and hosts transformed with said vectors.

(57) Novel carbonyl hydrolase mutants derived from the amino acid sequence of naturally-occurring or recombinant non-human carbonyl hydrolases and DNA sequences encoding the same. The mutant carbonyl hydrolases, in general, are obtained by *in vitro* modification of a precursor DNA sequence encoding the naturally-occurring or recombinant carbonyl hydrolase to encode the substitution, insertion or deletion of one or more amino acids in the amino acid sequence of a precursor carbonyl hydrolase. Such mutants have one or more properties which are different than the same property of the precursor hydrolase.

NON-HUMAN CARBONYL HYDROLASE MUTANTS,  
DNA SEQUENCES AND VECTORS ENCODING SAME  
AND HOSTS TRANSFORMED WITH SAID VECTORS

The recent development of various in vitro techniques to manipulate the DNA sequences encoding naturally-occurring polypeptides as well as recent developments in the chemical synthesis of relatively short sequences of single and double stranded DNA has resulted in the speculation that such techniques can be used to modify enzymes to improve some functional property in a predictable way. Ulmer, K.M. (1983) Science 219, 666-671. The only working example disclosed therein is the substitution of a single amino acid within the active site of tyrosyl-tRNA synthetase (Cys35-Ser) which lead to a reduction in enzymatic activity. See Winter, G., et al. (1982) Nature 299, 756-758; and Wilkinson, A.J., et al. (1983) Biochemistry 22, 3581-3586 (Cys35-Gly mutation also resulted in decreased activity).

When the same t-RNA synthetase was modified by substituting a different amino acid residue within the active site with two different amino acids, one of the mutants (Thr51-Ala) reportedly demonstrated a predicted moderate increase in  $k_{cat}/K_m$  whereas a second mutant (Thr51-Pro) demonstrated a massive increase in  $k_{cat}/K_m$  which could not be explained with

certainty. Wilkinson, A.H., et al. (1984) Nature 307, 187-188.

Another reported example of a single substitution of an amino acid residue is the substitution of cysteine for isoleucine at the third residue of T4 lysozyme. Perry, L.J., et al. (1984) Science 226, 555-557. The resultant mutant lysozyme was mildly oxidized to form a disulfide bond between the new cysteine residue at position 3 and the native cysteine at position 97. This crosslinked mutant was initially described by the author as being enzymatically identical to, but more thermally stable than, the wild type enzyme. However, in a "Note Added in Proof", the author indicated that the enhanced stability observed was probably due to a chemical modification of cysteine at residue 54 since the mutant lysozyme with a free thiol at Cys54 has a thermal stability identical to the wild type lysozyme.

Similarly, a modified dihydrofolate reductase from E.coli has been reported to be modified by similar methods to introduce a cysteine which could be crosslinked with a naturally-occurring cysteine in the reductase. Villafranca, D.E., et al. (1983) Science 222, 782-788. The author indicates that this mutant is fully reactive in the reduced state but has significantly diminished activity in the oxidized state. In addition, two other substitutions of specific amino acid residues are reported which resulted in mutants which had diminished or no activity.

EPO Publication No. 0130756 discloses the substitution of specific residues within B. amyloliquefaciens subtilisin with specific amino acids. Thus, Met222 has been substituted with all 19 other amino acids,

Gly166 with 9 different amino acids and Gly169 with Ala and Ser.

5 As set forth below, several laboratories have also reported the use of site directed mutagenesis to produce the mutation of more than one amino acid residue within a polypeptide.

10 The amino-terminal region of the signal peptide of the prolipoprotein of the E. coli outer membrane was stated to be altered by the substitution or deletion of residues 2 and 3 to produce a charge change in that region of the polypeptide. Inoyye, S., et al. (1982) Proc. Nat. Acad. Sci. USA 79, 3438-3441. The same  
15 laboratory also reported the substitution and deletion of amino acid residues 9 and 14 to determine the effects of such substitution on the hydrophobic region of the same signal sequence. Inouye, S., et al. (1984) J. Biol. Chem. 259, 3729-3733.

20 Double mutants in the active site of tyrosyl-t-RNA synthetase have also been reported. Carter, P.J., et al. (1984) Cell 38, 835-840. In this report, the improved affinity of the previously described Thr51+Pro mutant for ATP was probed by producing a  
25 second mutation in the active site of the enzyme. One of the double mutants, Gly35/Pro51, reportedly demonstrated an unexpected result in that it bound ATP in the transition state better than was expected from the two single mutants. Moreover, the author warns,  
30 at least for one double mutant, that it is not readily predictable how one substitution alters the effect caused by the other substitution and that care must be taken in interpreting such substitutions.

5 A mutant is disclosed in U.S. Patent No. 4,532,207,  
wherein a polyarginine tail was attached to the  
C-terminal residue of  $\beta$ -urogastrone by modifying the  
DNA sequence encoding the polypeptide. As disclosed,  
the polyarginine tail changed the electrophoretic  
mobility of the urogastrone-polyarginine hybrid  
permitting selective purification. The polyarginine  
was subsequently removed, according to the patentee,  
by a polyarginine specific exopeptidase to produce the  
purified urogastrone. Properly construed, this  
10 reference discloses hybrid polypeptides which do not  
constitute mutant polypeptides containing the  
substitution, insertion or deletion of one or more  
amino acids of a naturally occurring polypeptide.

15 Single and double mutants of rat pancreatic trypsin  
have also been reported. Craik, C.S., et al. (1985)  
Science 228, 291-297. As reported, glycine residues  
at positions 216 and 226 were replaced with alanine  
residues to produce three trypsin mutants (two single  
20 mutants and one double mutant). In the case of the  
single mutants, the authors stated expectation was to  
observe a differential effect on  $K_m$ . They instead  
reported a change in specificity ( $k_{cat}/K_m$ ) which was  
primarily the result of a decrease in  $k_{cat}$ . In  
25 contrast, the double mutant reportedly demonstrated a  
differential increase in  $K_m$  for lysyl and arginyl  
substrates as compared to wild type trypsin but had  
virtually no catalytic activity.

30 The references discussed above are provided solely for  
their disclosure prior to the filing date of the  
instant case, and nothing herein is to be construed as  
an admission that the inventors are not entitled to  
antedate such disclosure by virtue of prior invention  
35 or priority based on earlier filed applications.

Based on the above references, however, it is apparent that the modification of the amino acid sequence of wild type enzymes often results in the decrease or destruction of biological activity.

5 Accordingly, it is an object herein to provide carbonyl hydrolase mutants which have at least one property which is different from the same property of the carbonyl hydrolase precursor from which the amino acid of said mutant is derived.

10

It is a further object to provide mutant DNA sequences encoding such carbonyl hydrolase mutants as well as expression vectors containing such mutant DNA sequences.

15

Still further, another object of the present invention is to provide host cells transformed with such vectors as well as host cells which are capable of expressing such mutants either intracellularly or  
20 extracellularly.

25

30

35

Summary of the Invention

The invention includes carbonyl hydrolase mutants, preferably having at least one property which is substantially different from the same property of the precursor non-human carbonyl hydrolase from which the amino acid sequence of the mutant is derived. These properties include oxidative stability, substrate, specificity catalytic activity, thermal stability, alkaline stability, pH activity profile and resistance to proteolytic degradation. The precursor carbonyl hydrolase may be naturally occurring carbonyl hydrolases or recombinant carbonyl hydrolases. The amino acid sequence of the carbonyl hydrolase mutant is derived by the substitution, deletion or insertion of one or more amino acids of the precursor carbonyl hydrolase amino acid sequence.

The invention also includes mutant DNA sequences encoding such carbonyl hydrolase mutants. Further the invention includes expression vectors containing such mutant DNA sequences as well as host cells transformed with such vectors which are capable of expressing said carbonyl hydrolase mutants.

Brief Description of the Drawings

Figure 1 shows the nucleotide sequence of the coding strand, correlated with the amino acid sequence of B. amyloliquefaciens subtilisin gene. Promoter (p) ribosome binding site (rbs) and termination (term) regions of the DNA sequence as well as sequences encoding the presequence (PRE) putative prosequence (PRO) and mature form (MAT) of the hydrolase are also shown.

Figure 2 is a schematic diagram showing the substrate binding cleft of subtilisin together with substrate.

-7-

Figure 3 is a stereo view of the S-1 binding subsite of B. amyloliquefaciens subtilisin showing a lysine P-1 substrate bound in the site in two different ways. Figure 3A shows Lysine P-1 substrate bound to form a salt bridge with a Glu at position 156. Figure 3B shows Lysine P-1 substrate bound to form a salt bridge with Glu at position 166.

Figure 4 is a schematic diagram of the active site of subtilisin Asp32, His64 and Ser221.

Figures 5A and 5B depict the amino acid sequence of subtilisin obtained from various sources. The residues directly beneath each residue of B. amyloliquefaciens subtilisin are equivalent residues which (1) can be mutated in a similar manner to that described for B. amyloliquefaciens subtilisin, or (2) can be used as a replacement amino acid residue in B. amyloliquefaciens subtilisin. Figure 5C depicts conserved residues of B. amyloliquefaciens subtilisin when compared to other subtilisin sequences.

Figures 6A and 6B depict the inactivation of the mutants Met222L and Met222Q when exposed to various organic oxidants.

Figure 7 depicts the ultraviolet spectrum of Met222F subtilisin and the difference spectrum generated after inactivation by dodecanoic acid (DDCA).

Figure 8 shows the pattern of cyanogen bromide digests of untreated and DDCA oxidized subtilisin Met222F on high resolution SDS-pyridine peptide gels.



Figure 9 depicts a map of the cyanogen bromide fragments of Fig. 8 and their alignment with the sequence of subtilisin Met222F.

5 Figure 10 depicts the construction of mutations between codons 45 and 50 of B. amyloliquefaciens subtilisin.

10 Figure 11 depicts the construction of mutations between codons 122 and 127 of B. amyloliquefaciens subtilisin.

15 Figure 12 depicts the effect of DPDA on the activity of subtilisin mutants at positions 50 and 124 in subtilisin Met222F.

Figure 13 depicts the construction of mutations at codon 166 of B. amyloliquefaciens subtilisin.

20 Figure 14 depicts the effect of hydrophobicity of the P-1 substrate side-chain on the kinetic parameters of wild-type B. amyloliquefaciens subtilisin.

25 Figure 15 depicts the effect of position 166 side-chain substitutions on P-1 substrate specificity. Figure 15A shows position 166 mutant subtilisins containing non-branched alkyl and aromatic side-chain substitutions arranged in order of increasing molecular volume. Figure 15B shows a series of mutant enzymes progressing through  $\beta$ - and  $\gamma$ -branched  
30 aliphatic side chain substitutions of increasing molecular volume.

35 Figure 16 depicts the effect of position 166 side-chain volume on log  $k_{cat}/K_m$  for various P-1 substrates.

5 Figure 17 shows the substrate specificity differences between Ile166 and wild-type (Gly166) B. amyloliquefaciens subtilisin against a series of alphatic and aromatic substrates. Each bar represents the difference in log kcat/Km for Ile166 minus wild-type (Gly166) subtilisin.

Figure 18 depicts the construction of mutations at codon 169 of B. amyloliquefaciens subtilisin.

10 Figure 19 depicts the construction of mutations at codon 104 of B. amyloliquefaciens subtilisin.

15 Figure 20 depicts the construction of mutations at codon 152 B. amyloliquefaciens subtilisin.

Figure 21 depicts the construction of single mutations at codon 156 and double mutations at codons 156 and 166 of B. amyloliquefaciens subtilisin.

20 Figure 22 depicts the construction of mutations at codon 217 for B. amyloliquefaciens subtilisin.

25 Figure 23 depicts the kcat/Km versus pH profile for mutations at codon 156 and 166 in B. amyloliquefaciens subtilisin.

30 Figure 23A depicts the kcat/Km versus pH profile for mutations at codon 156 and 166 in B. amyloliquefaciens subtilisin.

Figure 24 depicts the kcat/Km versus pH profile for mutations at codon 222 in B. amyloliquefaciens subtilisin.

35

Figure 25 depicts the constructing mutants at codons 94, 95 and 96.

5      Figures 26 and 27 depict substrate specificity of various wild type and mutant subtilisins for different substrates.

10      Figures 28 A, B, C and D depict the effect of charge in the P-1 binding sites due to substitutions at codon 156 and 166.

15      Figures 29 A and B are a stereoview of the P-1 binding site of subtilisin BPN' showing a lysine P-1 substrate bound in the site in two ways. In 29A, Lysine P-1 substrate is built to form a salt bridge with a Glu at codon 156. In 29B, Lysine P-1 substrate is built to form a salt bridge with Glu at codon 166.

20      Figure 30 demonstrates residual enzyme activity versus temperature curves for purified wild-type (Panel A), C22/C87 (Panel B) and C24/C87 (Panel C).

25      Figure 31 depicts the strategy for producing point mutations in the subtilisin coding sequence by misincorporation of  $\alpha$ -thioldeoxynucleotide triphosphates.

30      Figure 32 depicts the autolytic stability of purified wild type and mutant subtilisins 170E, 107V, 213R and 107V/213R at alkaline pH.

35      Figure 33 depicts the autolytic stability of purified wild type and mutant subtilisins V50, F50 and F50/V107/R213 at alkaline pH.

Figure 34 depicts the strategy for constructing plasmids containing random cassette mutagenesis over residues 197 through 228.

5 Figure 35 depicts the oligodeoxynucleotides used for random cassette mutagenesis over residues 197 through 228.

10 Figure 36 depicts the construction of mutants at codon 204.

Figure 37 depicts the oligodeoxynucleotides used for synthesizing mutants at codon 204.

#### Detailed Description

15 The inventors have discovered that various single and multiple in vitro mutations involving the substitution, deletion or insertion of one or more amino acids within a non-human carbonyl hydrolase amino acid sequence can confer advantageous properties  
20 to such mutants when compared to the non-mutated carbonyl hydrolase.

Specifically, B. amyloliquefaciens subtilisin, an alkaline bacterial protease, has been mutated by  
25 modifying the DNA encoding the subtilisin to encode the substitution of one or more amino acids at various amino acid residues within the mature form of the subtilisin molecule. These in vitro mutant subtilisins have at least one property which is  
30 different when compared to the same property of the precursor subtilisin. These modified properties fall into several categories including: oxidative stability, substrate specificity, thermal stability, alkaline stability, catalytic activity, pH activity

35

-12-

profile, resistance to proteolytic degradation,  $K_m$ ,  $k_{cat}$  and  $K_m/k_{cat}$  ratio.

Carbonyl hydrolases are enzymes which hydrolyze

O

||

5

compounds containing C-X bonds in which X is oxygen or nitrogen. They include naturally-occurring carbonyl hydrolases and recombinant carbonyl hydrolases. Naturally occurring carbonyl hydrolases principally include hydrolases, e.g. lipases and peptide hydrolases, e.g. subtilisins or metalloproteases. Peptide hydrolases include  $\alpha$ -aminoacylpeptide hydrolase, peptidylamino-acid hydrolase, acylamino hydrolase, serine carboxypeptidase, metallocarboxypeptidase, thiol proteinase, carboxylproteinase and metalloproteinase. Serine, metallo, thiol and acid proteases are included, as well as endo and exo-proteases.

10

15

20

25

"Recombinant carbonyl hydrolase" refers to a carbonyl hydrolase in which the DNA sequence encoding the naturally occurring carbonyl hydrolase is modified to produce a mutant DNA sequence which encodes the substitution, insertion or deletion of one or more amino acids in the carbonyl hydrolase amino acid sequence. Suitable modification methods are disclosed herein and in EPO Publication No. 0130756 published January 9, 1985.

30

35

Subtilisins are bacterial carbonyl hydrolases which generally act to cleave peptide bonds of proteins or peptides. As used herein, "subtilisin" means a naturally occurring subtilisin or a recombinant subtilisin. A series of naturally occurring subtilisins is known to be produced and often secreted

by various bacterial species. Amino acid sequences of the members of this series are not entirely homologous. However, the subtilisins in this series exhibit the same or similar type of proteolytic activity. This class of serine proteases shares a common amino acid sequence defining a catalytic triad which distinguishes them from the chymotrypsin related class of serine proteases. The subtilisins and chymotrypsin related serine proteases both have a catalytic triad comprising aspartate, histidine and serine. In the subtilisin related proteases the relative order of these amino acids, reading from the amino to carboxy terminus is aspartate-histidine-serine. In the chymotrypsin related proteases the relative order, however is histidine-aspartate-serine. Thus, subtilisin herein refers to a serine protease having the catalytic triad of subtilisin related proteases.

"Recombinant subtilisin" refers to a subtilisin in which the DNA sequence encoding the subtilisin is modified to produce a mutant DNA sequence which encodes the substitution, deletion or insertion of one or more amino acids in the naturally occurring subtilisin amino acid sequence. Suitable methods to produce such modification include those disclosed herein and in EPO Publication No. 0130756. For example, the subtilisin multiple mutant herein containing the substitution of methionine at amino acid residues 50, 124 and 222 with phenylalanine, isoleucine and glutamine, respectively, can be considered to be derived from the recombinant subtilisin containing the substitution of glutamine at residue 222 (Q222) disclosed in EPO Publication No. 0130756. The multiple mutant thus is produced by the substitution of phenylalanine for methionine at

residue 50 and isoleucine for methionine at residue 124 in the Q222 recombinant subtilisin.

"Carbonyl hydrolases" and their genes may be obtained from many procaryotic and eucaryotic organisms. Suitable examples of procaryotic organisms include gram negative organisms such as E. coli or pseudomonas and gram positive bacteria such as micrococcus or bacillus. Examples of eucaryotic organisms from which carbonyl hydrolase and their genes may be obtained include yeast such as S. cerevisiae, fungi such as Aspergillus sp., and non-human mammalian sources such as, for example, Bovine sp. from which the gene encoding the carbonyl hydrolase chymosin can be obtained. As with subtilisins, a series of carbonyl hydrolases can be obtained from various related species which have amino acid sequences which are not entirely homologous between the members of that series but which nevertheless exhibit the same or similar type of biological activity. Thus, non-human carbonyl hydrolase as used herein has a functional definition which refers to carbonyl hydrolases which are associated, directly or indirectly, with procaryotic and non-human eucaryotic sources.

A "carbonyl hydrolase mutant" has an amino acid sequence which is derived from the amino acid sequence of a non-human "precursor carbonyl hydrolase". The precursor carbonyl hydrolases include naturally-occurring carbonyl hydrolases and recombinant carbonyl hydrolases. The amino acid sequence of the carbonyl hydrolase mutant is "derived" from the precursor hydrolase amino acid sequence by the substitution, deletion or insertion of one or more amino acids of the precursor amino acid sequence. Such modification is of the "precursor DNA sequence" which encodes the

amino acid sequence of the precursor carbonyl hydrolase rather than manipulation of the precursor carbonyl hydrolase per se. Suitable methods for such manipulation of the precursor DNA sequence include methods disclosed herein and in EPO Publication No. 0130756.

Specific residues of B. amyloliquefaciens subtilisin are identified for substitution, insertion or deletion. These amino acid position numbers refer to those assigned to the B. amyloliquefaciens subtilisin sequence presented in Fig. 1. The invention, however, is not limited to the mutation of this particular subtilisin but extends to precursor carbonyl hydrolases containing amino acid residues which are "equivalent" to the particular identified residues in B. amyloliquefaciens subtilisin.

A residue (amino acid) of a precursor carbonyl hydrolase is equivalent to a residue of B. amyloliquefaciens subtilisin if it is either homologous (i.e., corresponding in position in either primary or tertiary structure) or analogous to a specific residue or portion of that residue in B. amyloliquefaciens subtilisin (i.e., having the same or similar functional capacity to combine, react, or interact chemically).

In order to establish homology to primary structure, the amino acid sequence of a precursor carbonyl hydrolase is directly compared to the B. amyloliquefaciens subtilisin primary sequence and particularly to a set of residues known to be invariant in all subtilisins for which sequence is known (Figure 5C). After aligning the conserved residues, allowing for necessary insertions and



-16-

5 deletions in order to maintain alignment (i.e.,  
avoiding the elimination of conserved residues through  
arbitrary deletion and insertion), the residues  
equivalent to particular amino acids in the primary  
sequence of B. amyloliquefaciens subtilisin are  
defined. Alignment of conserved residues preferably  
should conserve 100% of such residues. However,  
alignment of greater than 75% or as little as 50% of  
conserved residues is also adequate to define  
10 equivalent residues. Conservation of the catalytic  
triad, Asp32/His64/Ser221 should be maintained.

15 For example, in Figure 5A the amino acid sequence of  
subtilisin from B. amyloliquefaciens B. subtilisin  
var. Il68 and B. lichenformis (carlsbergensis) are  
aligned to provide the maximum amount of homology  
between amino acid sequences. A comparison of these  
sequences shows that there are a number of conserved  
residues contained in each sequence. These residues  
are identified in Fig. 5C.

20 These conserved residues thus may be used to define  
the corresponding equivalent amino acid residues of B.  
amyloliquefaciens subtilisin in other carbonyl  
hydrolases such as thermitase derived from  
25 Thermoactinomyces. These two particular sequences are  
aligned in Fig. 5B to produce the maximum homology of  
conserved residues. As can be seen there are a number  
of insertions and deletions in the thermitase sequence  
as compared to B. amyloliquefaciens subtilisin. Thus,  
30 in thermitase the equivalent amino acid of Tyr217 in  
B. amyloliquefaciens subtilisin is the particular  
lysine shown beneath Tyr217.

35 In Fig. 5A, the equivalent amino acid at position 217  
in B. amyloliquefaciens subtilisin is Tyr. Likewise,

in B. subtilis subtilisin position 217 is also occupied by Tyr but in B. licheniformis position 217 is occupied by Leu.

5 Thus, these particular residues in thermitase, and subtilisin from B. subtilis and B. licheniformis may be substituted by a different amino acid to produce a mutant carbonyl hydrolase since they are equivalent in primary structure to Tyr217 in B. amyloliquefaciens subtilisin. Equivalent amino acids of course are not  
10 limited to those for Tyr217 but extend to any residue which is equivalent to a residue in B. amyloliquefaciens whether such residues are conserved or not.

15 Equivalent residues homologous at the level of tertiary structure for a precursor carbonyl hydrolase whose tertiary structure has been determined by x-ray crystallography, are defined as those for which the atomic coordinates of 2 or more of the main chain atoms of a particular amino acid residue of the  
20 precursor carbonyl hydrolase and B. amyloliquefaciens subtilisin (N on N, CA on CA, C on C, and O on O) are within 0.13nm and preferably 0.1nm after alignment. Alignment is achieved after the best model has been oriented and positioned to give the maximum overlap of  
25 atomic coordinates of non-hydrogen protein atoms of the carbonyl hydrolase in question to the B. amyloliquefaciens subtilisin. The best model is the crystallographic model giving the lowest R factor for experimental diffraction data at the highest  
30 resolution available.

$$R \text{ factor} = \frac{\sum_h |F_o(h)| - |F_c(h)|}{\sum_h |F_o(h)|}$$

Equivalent residues which are functionally analogous to a specific residue of B. amyloliquefaciens subtilisin are defined as those amino acids of the precursor carbonyl hydrolases which may adopt a conformation such that they either alter, modify or contribute to protein structure, substrate binding or catalysis in a manner defined and attributed to a specific residue of the B. amyloliquefaciens subtilisin as described herein. Further, they are those residues of the precursor carbonyl hydrolase (for which a tertiary structure has been obtained by x-ray crystallography), which occupy an analogous position to the extent that although the main chain atoms of the given residue may not satisfy the criteria of equivalence on the basis of occupying a homologous position, the atomic coordinates of at least two of the side chain atoms of the residue lie within 0.13nm of the corresponding side chain atoms of B. amyloliquefaciens subtilisin. The three dimensional structures would be aligned as outlined above.

Some of the residues identified for substitution, insertion or deletion are conserved residues whereas others are not. In the case of residues which are not conserved, the replacement of one or more amino acids is limited to substitutions which produce a mutant which has an amino acid sequence that does not correspond to one found in nature. In the case of conserved residues, such replacements should not result in a naturally occurring sequence. The carbonyl hydrolase mutants of the present invention include the mature forms of carbonyl hydrolase mutants as well as the pro- and prepro-forms of such hydrolase mutants. The prepro-forms are the preferred construction since

this facilitates the expression, secretion and maturation of the carbonyl hydrolase mutants.

5 "Expression vector" refers to a DNA construct containing a DNA sequence which is operably linked to a suitable control sequence capable of effecting the expression of said DNA in a suitable host. Such control sequences include a promoter to effect transcription, an optional operator sequence to control such transcription, a sequence encoding suitable mRNA ribosome binding sites, and sequences 10 which control termination of transcription and translation. The vector may be a plasmid, a phage particle, or simply a potential genomic insert. Once transformed into a suitable host, the vector may replicate and function independently of the host genome, or may, in some instances, integrate into the 15 genome itself. In the present specification, "plasmid" and "vector" are sometimes used interchangeably as the plasmid is the most commonly used form of vector at present. However, the 20 invention is intended to include such other forms of expression vectors which serve equivalent functions and which are, or become, known in the art.

25 The "host cells" used in the present invention generally are procaryotic or eucaryotic hosts which preferably have been manipulated by the methods disclosed in EPO Publication No. 0130756 to render

them incapable of secreting enzymatically active endoprotease. A preferred host cell for expressing subtilisin is the *Bacillus* strain BG2036 which is deficient in enzymatically active neutral protease and alkaline protease (subtilisin). The construction of strain BG2036 is described in detail in EPO Publication No. 0130756 and further described by Yang, M.Y., et al. (1984) J. Bacteriol. 160, 15-21. Other host cells for expressing subtilisin include Bacillus subtilis 1168 (EPO Publication No. 0130756).

Host cells are transformed or transfected with vectors constructed using recombinant DNA techniques. Such transformed host cells are capable of either replicating vectors encoding the carbonyl hydrolase mutants or expressing the desired carbonyl hydrolase mutant. In the case of vectors which encode the pre or prepro form of the carbonyl hydrolase mutant, such mutants, when expressed, are typically secreted from the host cell into the host cell medium.

"Operably linked" when describing the relationship between two DNA regions simply means that they are functionally related to each other. For example, a presequence is operably linked to a peptide if it functions as a signal sequence, participating in the secretion of the mature form of the protein most probably involving cleavage of the signal sequence. A promoter is operably linked to a coding sequence if it controls the transcription of the sequence; a ribosome binding site is operably linked to a coding sequence if it is positioned so as to permit translation.

The genes encoding the naturally-occurring precursor carbonyl hydrolase may be obtained in accord with the

general methods described herein in EPO Publication No. 0130756.

5 Once the carbonyl hydrolase gene has been cloned, a number of modifications are undertaken to enhance the use of the gene beyond synthesis of the naturally-  
10 occurring precursor carbonyl hydrolase. Such modifications include the production of recombinant carbonyl hydrolases as disclosed in EPO Publication No. 0130756 and the production of carbonyl hydrolase mutants described herein.

The carbonyl hydrolase mutants of the present invention may be generated by site specific mutagenesis (Smith, M. (1985) Ann. Rev. Genet. 423;  
15 Zoeller, M.J., et al. (1982) Nucleic Acid Res. 10, 6487-6500), cassette mutagenesis (EPO Publication No. 0130756) or random mutagenesis (Shortle, D., et al. (1985) Genetics, 110, 539; Shortle, D., et al. (1986) Proteins: Structure, Function and Genetics, 1, 81;  
20 Shortle, D. (1986) J. Cell. Biochem., 30, 281; Alber, T., et al. (1985) Proc. Natl. Acad. of Sci., 82, 747; Matsumura, M., et al. (1985) J. Biochem., 260, 15298; Liao, H., et al. (1986) Proc. Natl. Acad. of Sci., 83 576) of the cloned precursor carbonyl hydrolase.  
25 Cassette mutagenesis and the random mutagenesis method disclosed herein are preferred.

The mutant carbonyl hydrolases expressed upon transformation of suitable hosts are screened for  
30 enzymes exhibiting one or more properties which are substantially different from the properties of the precursor carbonyl hydrolases, e.g., changes in substrate specificity, oxidative stability, thermal stability, alkaline stability, resistance to

35

proteolytic degradation, pH-activity profiles and the like.

5 A change in substrate specificity is defined as a difference between the  $k_{cat}/K_m$  ratio of the precursor carbonyl hydrolase and that of the hydrolase mutant. The  $k_{cat}/K_m$  ratio is a measure of catalytic efficiency. Carbonyl hydrolase mutants with increased or diminished  $k_{cat}/K_m$  ratios are described in the examples. Generally, the objective will be to secure  
10 a mutant having a greater (numerically large)  $k_{cat}/K_m$  ratio for a given substrate, thereby enabling the use of the enzyme to more efficiently act on a target substrate. A substantial change in  $k_{cat}/K_m$  ratio is preferably at least 2-fold increase or decrease.  
15 However, smaller increases or decreases in the ratio (e.g., at least 1.5-fold) are also considered substantial. An increase in  $k_{cat}/K_m$  ratio for one substrate may be accompanied by a reduction in  $k_{cat}/K_m$  ratio for another substrate. This is a shift in  
20 substrate specificity, and mutants exhibiting such shifts have utility where the precursor hydrolase is undesirable, e.g. to prevent undesired hydrolysis of a particular substrate in an admixture of substrates.  $K_m$  and  $k_{cat}$  are measured in accord with known  
25 procedures, as described in EPO Publication No. 0130756 or as described herein.

Oxidative stability is measured either by known procedures or by the methods described hereinafter. A  
30 substantial change in oxidative stability is evidenced by at least about 50% increase or decrease (preferably decrease) in the rate of loss of enzyme activity when exposed to various oxidizing conditions. Such oxidizing conditions are exposure to the organic  
35

-23-

oxidant dimerdodecanoic acid (DPDA) under the conditions described in the examples.

Alkaline stability is measured either by known procedures or by the methods described herein. A  
5 substantial change in alkaline stability is evidenced by at least about a 5% or greater increase or decrease (preferably increase) in the half life of the enzymatic activity of a mutant when compared to the precursor carbonyl hydrolase. In the case of  
10 subtilisins, alkaline stability was measured as a function of autoproteolytic degradation of subtilisin at alkaline pH, e.g. for example, 0.1M sodium phosphate, pH 12 at 25° or 30°C.

15 Thermal stability is measured either by known procedures or by the methods described herein. A substantial change in thermal stability is evidenced by at least about a 5% or greater increase or decrease (preferably increase) in the half-life of the  
20 catalytic activity of a mutant when exposed to a relatively high temperature and neutral pH as compared to the precursor carbonyl hydrolase. In the case of subtilisins, thermal stability is measured by the autoproteolytic degradation of subtilisin at elevated  
25 temperatures and neutral pH, e.g., for example 2mM calcium chloride, 50mM MOPS pH 7.0 at 59°C.

The inventors have produced mutant subtilisins containing the substitution of the amino acid residues  
30 of B. amyloliquefaciens subtilisin shown in Table I. The wild type amino acid sequence and DNA sequence of B. amyloliquefaciens subtilisin is shown in Fig. 1.

35



TABLE I

<u>Residue</u>		<u>Replacement Amino Acid</u>
	Tyr21	F A
	Thr22	C
	Ser24	C
5	Asp32	Q S
	Ser33	A T
	Asp36	A G
	Gly46	V
	Ala48	E V R
10	Ser49	C L
	Met50	C F V
	Asn77	D
	Ser87	C
	Lys94	C
15	Val95	C
	Leu96	D
	Tyr104	A C D E F G H I K L M N P Q R S T V W
	Ile107	V
	Gly110	C R
20	Met124	I L
	Asn155	A D H Q T
	Glu156	Q S
	Gly166	C E I L M P S T W Y
	Gly169	C D E F H I K L M N P Q R T V W Y
25	Lys170	E R
	Tyr171	F
	Pro172	E Q
	Phe189	A C D E G H I K L M N P Q R S T V W Y
	Asp197	R A
30	Met199	I
	Ser204	C R L P
	Lys213	R T
	Tyr217	A C D E F G H I K L M N P Q R S T V W
35	Ser221	A C

The different amino acids substituted are represented in Table I by the following single letter designations:

	<u>Amino acid or residue thereof</u>	<u>3-letter symbol</u>	<u>1-letter symbol</u>
5	Alanine	Ala	A
	Glutamate	Glu	E
	Glutamine	Gln	Q
10	Aspartate	Asp	D
	Asparagine	Asn	N
	Leucine	Leu	L
	Glycine	Gly	G
	Lysine	Lys	K
15	Serine	Ser	S
	Valine	Val	V
	Arginine	Arg	R
	Threonine	Thr	T
	Proline	Pro	P
20	Isoleucine	Ile	I
	Methionine	Met	M
	Phenylalanine	Phe	F
	Tyrosine	Tyr	Y
	Cysteine	Cys	C
25	Tryptophan	Trp	W
	Histidine	His	H

Except where otherwise indicated by context, wild-type amino acids are represented by the above three-letter symbols and replaced amino acids by the above single-letter symbols. Thus, if the methionine at residue 50 in *B. amyloliquefaciens* subtilisin is

replaced by phenylalanine, this mutation (mutant) may be designated Met50F or F50. Similar designations are used for multiple mutants.

5 In addition to the amino acids used to replace the residues disclosed in Table I, other replacements of amino acids at these residues are expected to produce mutant subtilisins having useful properties. These residues and replacement amino acids are shown in Table II.

10

15

20

25

30

35

TABLE II

	<u>Residue</u>	<u>Replacement Amino Acid(s)</u>
	Tyr-21	L
	Thr22	K
5	Ser24	A
	Asp32	
	Ser33	G
	Gly46	
	Ala48	
10	Ser49	
	Met50	L K I V
	Asn77	D
	Ser87	N
	Lys94	R Q
15	Val95	L I
	Tyr104	
	Met124	K A
	Ala152	C L I T M
	Asn155	
20	Glu156	A T M L Y
	Gly166	
	Gly169	
	Tyr171	K R E Q
	Prol72	D N
25	Phel89	
	Tyr217	
	Ser221	
	Met222	

30

Each of the mutant subtilisins in Table I contain the replacement of a single residue of the B. amyloliquefaciens amino acid sequence. These particular residues were chosen to probe the influence

35

of such substitutions on various properties of B. amyloliquefacien subtilisin.

Thus, the inventors have identified Met124 and Met222 as important residues which if substituted with another amino acid produce a mutant subtilisin with enhanced oxidative stability. For Met124, Leu and Ile are preferred replacement amino acids. Preferred amino acids for replacement of Met222 are disclosed in EPO Publication No. 0130756.

Various other specific residues have also been identified as being important with regard to substrate specificity. These residues include Tyr104, Ala152, Glu156, Gly166, Gly169, Phe189 and Tyr217 for which mutants containing the various replacement amino acids presented in Table I have already been made, as well as other residues presented below for which mutants have yet to be made.

The identification of these residues, including those yet to be mutated, is based on the inventors' high resolution crystal structure of B. amyloliquefaciens subtilisin to 1.8 Å (see Table III), their experience with in vitro mutagenesis of subtilisin and the literature on subtilisin. This work and the x-ray crystal structures of subtilisin containing covalently bound peptide inhibitors (Robertus, J.D., et al. (1972) Biochemistry 11, 2439-2449), product complexes (Robertus, J.D., et al. (1972) Biochemistry 11, 4293-4303), and transition state analogs (Matthews, D.A., et al. (1975) J. Biol. Chem. 250, 7120-7126; Poulos, T.L., et al. (1976) J. Biol. Chem. 251, 1097-1103), has helped in identifying an extended peptide binding cleft in subtilisin. This substrate binding cleft together with substrate is schematically

diagrammed in Fig. 2, according to the nomenclature of Schechter, I., et al. (1967) Biochem Bio. Res. Commun. 27, 157. The scissile bond in the substrate is identified by an arrow. The P and P' designations refer to the amino acids which are positioned  
5 respectively toward the amino or carboxy terminus relative to the scissile bond. The S and S' designations refer to subsites in the substrate binding cleft of subtilisin which interact with the corresponding substrate amino acid residues.

10

15

20

25

30

35

-30-

Atomic Coordinates for the  
Apoenzyme Form of B. Amyloliquefaciens  
Subtilisin to 1.8Å Resolution

1	ALA N	19.434	53.195	-21.756	1	ALA CA	19.811	51.774	-21.965
1	ALA C	18.731	50.995	-21.324	1	ALA O	18.376	51.197	-20.175
1	ALA CB	21.099	51.318	-21.183	2	GLN N	18.268	49.886	-22.041
2	GLN CA	17.219	49.008	-21.434	2	GLN C	17.875	47.706	-20.992
2	GLN O	18.765	47.165	-21.691	2	GLN CB	16.125	48.760	-22.449
2	GLN CG	15.328	47.905	-21.927	2	GLN CO	13.912	47.762	-22.930
2	GLN OE1	13.023	49.612	-22.867	2	GLN NE2	14.115	46.917	-23.926
3	SER N	17.477	47.205	-19.852	3	SER CA	17.950	45.868	-19.437
3	SER C	16.735	44.918	-19.490	3	SER O	15.590	45.352	-19.229
3	SER CB	18.588	45.838	-18.069	3	SER OG	17.482	46.210	-17.049
4	VAL N	16.991	43.444	-19.725	4	VAL CA	15.946	42.619	-19.639
4	VAL C	16.129	41.934	-18.290	4	VAL O	17.123	41.178	-18.086
4	VAL CB	16.008	41.622	-20.822	4	VAL CG1	14.874	40.572	-20.741
4	VAL CG2	16.037	42.266	-22.186	5	PRO N	15.239	42.104	-17.331
5	PRO CA	15.384	41.415	-16.027	5	PRO C	15.501	39.905	-16.249
5	PRO O	14.885	39.263	-17.146	5	PRO CB	14.150	41.880	-15.263
5	PRO CG	13.841	43.215	-15.921	5	PRO CO	14.844	42.986	-17.417
6	TYR N	16.363	39.240	-15.487	6	TYR CA	16.628	37.803	-15.715
6	TYR C	15.359	36.975	-15.528	6	TYR O	15.224	35.943	-16.235
6	TYR CB	17.824	37.323	-14.834	6	TYR CG	18.021	35.847	-15.055
6	TYR CD1	18.437	35.452	-16.346	6	TYR CD2	17.696	34.908	-14.071
6	TYR CE1	18.535	34.970	-16.653	6	TYR CE2	17.815	33.539	-14.379
6	TYR CZ	18.222	33.154	-15.621	6	TYR OH	18.312	31.838	-15.996
7	GLY N	14.464	37.362	-14.630	7	GLY CA	13.211	36.640	-14.376
7	GLY C	12.600	36.535	-15.670	7	GLY O	11.747	35.478	-15.883
8	VAL N	12.441	37.529	-16.541	8	VAL CA	11.777	37.523	-17.836
8	VAL C	12.363	36.433	-18.735	8	VAL O	11.639	35.716	-19.470
8	VAL CB	11.765	38.900	-18.567	8	VAL CG1	11.106	38.893	-19.943
8	VAL CG2	18.991	39.919	-17.733	9	SER N	13.661	36.318	-18.775
9	SER CA	14.419	35.362	-19.562	9	SER C	14.188	33.920	-18.965
9	SER O	14.112	33.014	-19.901	9	SER CB	15.926	35.632	-19.505
9	SER OG	16.162	36.747	-20.358	10	GLN N	14.115	33.887	-17.662
10	GLN CA	13.964	32.636	-16.876	10	GLN C	12.687	31.887	-17.277
10	GLN O	12.785	30.642	-17.413	10	GLN CB	14.125	32.885	-15.410
10	GLN CG	14.295	31.617	-14.588	10	GLN CO	14.486	31.911	-13.147
10	GLN OE1	14.554	33.868	-12.744	10	GLN NE2	14.552	30.960	-12.251
11	ILE N	11.625	32.575	-17.678	11	ILE CA	10.373	31.904	-18.182
11	ILE C	10.209	31.792	-19.605	11	ILE O	9.173	31.333	-20.180
11	ILE CB	9.132	32.669	-17.675	11	ILE CG1	9.046	34.117	-18.049
11	ILE CG2	9.162	32.655	-15.941	11	ILE CO1	7.588	34.648	-17.923
12	LYS N	11.272	32.185	-20.277	12	LYS CA	11.388	32.119	-21.722
12	LYS C	10.456	33.886	-22.522	12	LYS O	10.178	32.703	-23.686
12	LYS CB	11.257	30.446	-22.216	12	LYS CG	12.283	29.830	-21.423
12	LYS CO	12.543	28.517	-22.159	12	LYS CE	13.023	27.467	-21.166
12	LYS NZ	14.476	27.680	-20.935	13	ALA N	18.109	34.138	-21.991
13	ALA CA	9.325	35.198	-22.631	13	ALA C	18.026	35.716	-23.863
13	ALA O	9.338	35.804	-24.981	13	ALA CB	8.885	36.195	-21.565
14	PRO N	11.332	35.950	-23.893	14	PRO CA	11.985	36.430	-25.120
14	PRO C	11.786	35.557	-26.317	14	PRO O	11.778	36.047	-27.445
14	PRO CB	13.462	36.580	-24.692	14	PRO CG	13.328	36.970	-23.221
14	PRO CO	12.281	35.936	-22.758	15	ALA N	11.560	34.236	-26.129
15	ALA CA	11.379	33.458	-27.367	15	ALA C	10.082	33.795	-28.032
15	ALA O	10.008	33.718	-29.278	15	ALA CB	11.592	31.969	-27.062
16	LEU N	9.085	34.138	-27.240	16	LEU CA	7.791	34.558	-27.828
16	LEU C	7.912	35.925	-28.521	16	LEU O	7.342	36.126	-29.588
16	LEU CB	6.746	34.623	-26.698	16	LEU CG	5.790	33.465	-26.522
16	LEU CD1	5.881	33.234	-27.809	16	LEU CD2	6.694	32.207	-26.283
17	WIS N	8.665	36.828	-27.922	17	WIS CA	8.890	38.151	-28.539
17	WIS C	9.510	37.981	-29.898	17	WIS O	9.187	38.622	-30.856
17	WIS CB	9.708	39.180	-27.652	17	WIS CG	9.185	39.288	-26.262
17	WIS ND1	9.930	39.837	-25.272	17	WIS CD2	8.088	38.924	-25.694
17	WIS CE1	9.226	39.914	-24.144	17	WIS NE2	8.079	39.328	-24.381
18	SER N	18.443	37.833	-38.822	18	SER CA	11.189	36.739	-31.322

10	SLY C	30.139	36.123	-32.343	10	SLY D	30.947	36.112	-33.534
10	SLY CD	12.311	33.799	-31.172	10	SLY DS	13.321	36.488	-30.399
10	SLM N	9.880	35.495	-31.943	10	SLM CA	0.082	36.961	-32.878
10	SLM C	7.142	36.111	-33.303	10	SLM D	6.297	35.972	-34.219
10	SLM CB	7.323	33.849	-32.280	10	SLM CG	7.975	32.802	-31.823
10	SLM CD	6.923	31.707	-31.181	10	SLM CCI	5.719	31.833	-31.444
10	SLM MIZ	7.362	30.852	-30.254	10	SLY N	7.285	37.223	-32.587
20	SLY CA	6.369	30.387	-32.859	20	SLY C	5.181	38.492	-31.880
20	SLY D	4.263	39.276	-32.215	20	SLY C	5.202	37.801	-30.761
21	TYR CA	4.118	37.831	-29.763	21	TYR N	4.579	38.552	-28.925
21	TYR D	5.422	38.074	-27.756	21	TYR C	3.498	36.431	-29.643
21	TYR CG	2.973	31.784	-30.709	21	TYR CCI	1.793	36.332	-31.238
21	TYR CD2	3.450	34.794	-31.397	21	TYR CCI	1.306	35.797	-32.446
21	TYR CCI	3.193	34.261	-32.588	21	TYR CCI	2.053	34.755	-33.047
21	TYR CM	1.301	34.241	-34.250	21	TYR CCI	3.902	39.090	-28.288
22	TYR CA	4.262	40.327	-27.129	22	TYR C	3.091	40.822	-26.244
22	TYR D	3.287	41.725	-25.325	22	TYR CCI	3.133	41.759	-27.611
22	TYR DGI	4.319	42.437	-25.597	22	TYR CCI	6.474	41.323	-28.229
23	GLY N	1.939	40.285	-26.453	23	GLY CA	0.809	40.600	-23.542
23	GLY C	-0.157	41.631	-26.118	23	GLY D	-1.013	42.095	-23.350
24	SLY N	-0.023	41.967	-27.371	24	SLY CA	-0.897	42.937	-28.912
24	SLY C	-2.383	42.626	-27.864	24	SLY D	-2.813	41.588	-28.168
24	SLY CB	-0.734	43.120	-29.320	24	SLY DG	0.543	43.432	-29.728
25	ASH N	-3.059	43.692	-27.515	25	ASH CA	-4.519	43.487	-27.393
25	ASH C	-9.013	42.973	-26.205	25	ASH D	-6.233	42.688	-28.190
25	ASH CB	-5.165	43.217	-28.700	25	ASH CG	-6.960	44.170	-29.883
25	ASH CCI	-4.969	43.747	-31.083	25	ASH CCI	-4.747	45.461	-29.594
26	VAL N	-4.177	42.449	-25.292	26	VAL CA	-4.674	41.679	-24.143
26	VAL C	-4.792	42.652	-22.917	26	VAL D	-3.858	43.419	-22.689
26	VAL CB	-3.714	40.503	-23.821	26	VAL CCI	-4.160	39.802	-22.548
26	VAL CCI	-3.598	39.576	-23.018	27	LVS N	-5.910	42.613	-21.301
27	LVS CA	-6.133	43.524	-21.175	27	LVS C	-5.815	42.872	-19.541
27	LVS D	-6.405	41.873	-19.413	27	LVS CB	-7.590	43.981	-21.149
27	LVS CG	-8.044	44.575	-22.490	27	LVS CCI	-9.321	45.302	-22.820
27	LVS CCI	-10.304	45.497	-23.137	27	LVS MZ	-9.886	46.253	-24.244
28	VAL N	-4.818	43.462	-19.200	28	VAL CA	-4.457	42.950	-17.897
28	VAL C	-4.758	43.959	-16.828	28	VAL D	-4.209	45.895	-16.817
28	VAL CB	-2.924	42.666	-17.032	28	VAL CCI	-2.466	42.193	-16.589
28	VAL CCI	-2.467	41.805	-19.173	29	ALA N	-5.484	43.527	-15.813
29	ALA CA	-5.747	44.330	-14.639	29	ALA C	-4.750	44.810	-13.552
29	ALA D	-4.664	42.845	-13.104	29	ALA CB	-7.172	44.187	-14.181
30	VAL N	-4.857	45.833	-13.072	30	VAL CA	-3.166	44.982	-11.910
30	VAL C	-3.958	45.409	-10.681	30	VAL D	-4.193	46.648	-18.878
30	VAL CB	-1.886	45.910	-12.149	30	VAL CCI	-0.906	45.901	-18.988
30	VAL CCI	-1.853	45.236	-13.307	31	LLE N	-4.514	44.315	-9.877
31	LLE CA	-5.328	44.846	-8.679	31	LLE C	-4.346	44.953	-7.846
31	LLE D	-3.825	43.913	-8.997	31	LLE CB	-4.457	43.776	-8.951
31	LLE CCI	-7.298	43.707	-9.798	31	LLE CCI	-7.278	40.838	-7.223
31	LLE CCI	-8.617	42.856	-9.717	32	ASP N	-4.044	46.193	-7.227
32	ASP CA	-2.944	46.467	-6.255	32	ASP C	-3.071	47.889	-5.783
32	ASP C	-4.197	48.418	-5.302	32	ASP CB	-1.695	46.129	-7.092
32	ASP CG	-8.483	45.702	-6.273	32	ASP CCI	0.834	44.592	-6.576
32	ASP CCI	-8.081	46.419	-5.330	33	SLY N	-1.931	48.812	-3.394
33	SLY CA	-1.895	49.857	-4.801	33	SLY C	-1.982	50.976	-3.888
33	SLY D	-1.706	52.136	-5.363	33	SLY CB	-0.621	49.922	-3.939
33	SLY DG	0.533	50.025	-4.774	34	GLY N	-2.173	50.740	-7.084
34	GLY CA	-2.155	51.728	-8.163	34	GLY C	-1.035	51.648	-9.057
34	GLY D	-0.144	50.831	-8.761	35	LLE N	-0.963	52.471	-10.102
35	LLE CA	0.208	52.438	-10.995	35	LLE C	0.548	53.919	-11.243
35	LLE D	-0.327	54.638	-11.744	35	LLE CB	-0.842	51.604	-12.367
35	LLE CCI	-0.530	50.210	-12.097	35	LLE CCI	1.149	51.741	-13.362
35	LLE CCI	-0.962	49.485	-13.424	36	ASP N	1.816	54.253	-10.971
36	ASP CA	2.359	55.618	-11.232	36	ASP C	2.281	55.956	-12.782



36 ASP D	3.004	55.471	-13.579	36 ASP CB	3.712	55.720	-10.514
36 ASP CG	4.339	57.099	-10.804	36 ASP DD1	3.755	57.974	-11.429
36 ASP DD2	5.449	57.277	-10.243	37 SER M	1.304	56.822	-13.111
37 SER CA	1.183	57.221	-14.512	37 SER C	2.377	58.095	-14.949
37 SER D	2.545	58.303	-16.151	37 SER CB	-0.093	58.069	-14.788
37 SER DG	-0.080	59.133	-13.879	38 SER M	3.163	58.614	-14.001
38 SER CA	4.261	59.505	-14.487	38 SER C	5.466	58.705	-14.992
38 SER D	6.543	59.251	-15.285	38 SER CB	4.742	60.435	-13.398
38 SER DG	5.376	59.865	-12.234	39 MIS M	5.454	57.390	-14.892
39 MIS CA	6.637	56.574	-15.291	39 MIS C	6.681	56.401	-16.778
39 MIS D	5.738	55.878	-17.419	39 MIS CB	6.637	55.203	-14.515
39 MIS CG	8.014	56.609	-14.456	39 MIS DD1	8.795	54.356	-15.561
39 MIS DD2	8.749	54.345	-13.389	39 MIS CE1	9.970	53.930	-15.130
39 MIS NE2	9.986	53.910	-13.800	40 PRD M	7.807	56.834	-17.387
40 PRD CA	7.988	56.697	-18.831	40 PRD C	8.154	55.280	-19.357
40 PRD D	8.832	55.097	-20.578	40 PRD CG	9.247	57.533	-19.161
40 PRD CG	10.853	57.485	-17.982	40 PRD CD	8.988	57.452	-16.776
41 ASP M	8.481	54.328	-18.495	41 ASP DD2	11.148	50.399	-18.668
41 ASP DD1	10.325	51.395	-20.429	41 ASP CG	10.473	51.307	-19.211
41 ASP CB	9.799	52.239	-18.224	41 ASP CA	8.645	52.959	-18.966
41 ASP C	7.311	52.163	-18.439	41 ASP D	7.396	50.947	-18.977
42 LEU M	6.185	52.803	-18.558	42 LEU CA	4.892	52.147	-18.466
42 LEU C	3.924	52.907	-19.376	42 LEU D	3.993	54.163	-19.490
42 LEU CB	4.421	52.158	-17.808	42 LEU CG	5.182	51.363	-15.946
42 LEU CD1	4.535	51.546	-14.581	42 LEU CD2	5.273	49.877	-16.350
43 LYS M	3.018	52.135	-19.944	43 LYS CA	1.893	52.685	-20.721
43 LYS C	0.637	52.156	-20.018	43 LYS D	0.504	58.920	-19.820
43 LYS CB	2.021	52.389	-22.169	43 LYS CG	0.685	52.436	-22.910
43 LYS CD	0.998	52.862	-24.339	43 LYS CE	-0.180	52.504	-25.260
43 LYS H2	0.337	51.757	-26.418	44 VAL M	-0.191	53.035	-19.490
44 VAL CA	-1.487	52.639	-18.765	44 VAL C	-2.571	52.887	-19.731
44 VAL D	-2.623	53.906	-20.434	44 VAL CB	-1.480	53.351	-17.383
44 VAL CG1	-2.724	52.941	-16.582	44 VAL CG2	-0.197	53.194	-16.553
45 ALA M	-3.494	51.951	-19.871	45 ALA CA	-4.619	51.977	-20.810
45 ALA C	-5.841	52.507	-20.053	45 ALA D	-6.783	53.085	-20.703
45 ALA CB	-4.831	50.580	-21.389	46 GLY M	-5.918	52.356	-18.748
46 GLY CA	-7.882	52.837	-18.001	46 GLY C	-6.987	52.443	-16.538
46 GLY D	-5.938	52.006	-16.035	47 GLY M	-8.892	52.658	-15.793
47 GLY CA	-8.014	52.246	-14.388	47 GLY C	-9.179	52.757	-13.572
47 GLY D	-9.988	53.481	-14.185	48 ALA M	-9.221	52.446	-12.330
48 ALA CA	-10.235	52.970	-11.382	48 ALA C	-9.790	52.675	-9.968
48 ALA D	-9.846	53.720	-9.725	48 ALA CB	-11.558	52.100	-11.617
49 SER M	-18.149	53.547	-9.837	49 SER CA	-9.752	53.355	-7.652
49 SER C	-10.947	52.986	-6.783	49 SER D	-11.972	53.677	-6.988
49 SER CB	-9.092	54.588	-7.029	49 SER DG	-8.879	54.255	-5.650
50 MET M	-10.835	52.007	-5.932	50 MET CA	-11.852	51.549	-4.974
50 MET C	-11.463	51.962	-3.561	50 MET D	-11.997	51.398	-2.575
50 MET CB	-12.012	50.818	-4.996	50 MET CG	-11.912	49.463	-6.389
50 MET SD	-13.468	49.889	-7.256	50 MET CE	-12.808	50.111	-0.903
51 VAL M	-10.427	52.740	-3.422	51 VAL CA	-9.968	53.170	-2.867
51 VAL C	-10.630	54.562	-1.987	51 VAL D	-10.237	55.437	-2.682
51 VAL CB	-8.443	53.155	-2.000	51 VAL CG1	-7.892	53.579	-0.631
51 VAL CG2	-7.764	53.815	-2.302	52 PRD M	-11.621	54.693	-1.056
52 PRD CA	-12.372	55.933	-0.821	52 PRD C	-11.490	57.123	-0.440
52 PRD D	-11.771	58.220	-0.925	52 PRD CB	-13.680	55.594	0.244
52 PRD CG	-13.583	54.103	0.085	52 PRD CD	-12.164	53.620	-0.175
53 SER M	-10.442	56.906	0.299	53 SER CA	-9.538	57.982	0.682
53 SER C	-8.428	58.245	-0.326	53 SER D	-7.679	59.224	-0.038
53 SER CB	-9.884	57.707	2.069	53 SER DG	-8.256	56.521	2.127
54 GLU M	-8.254	57.523	-1.393	54 GLU CA	-7.204	57.648	-2.621
54 GLU C	-7.747	57.303	-3.785	54 GLU D	-7.533	56.243	-4.379
54 GLU CB	-6.134	56.599	-2.154	54 GLU CG	-5.289	56.959	-0.927
54 GLU PD	-4.844	44.847	-0.078	54 GLU DFI	-3.445	56.404	-1.048

54	ELW DEZ	-3.908	55.777	0.271	55	TMR H	-0.971	58.251	-4.249
55	TMR CA	-9.433	58.121	-5.441	55	TMR C	-8.764	58.139	-6.779
55	TMR D	-9.433	57.919	-7.010	55	TMR CB	-10.586	59.200	-3.383
55	TMR DG1	-9.885	60.510	-5.418	55	TMR CG2	-11.432	59.143	-6.017
56	ASN H	-7.682	58.403	-6.877	56	ASN HD2	-4.930	61.179	-9.881
56	ASN DD1	-5.875	58.967	-10.337	56	ASN CG	-5.273	59.925	-0.555
56	ASN CB	-5.898	59.894	-8.208	56	ASN CA	-6.762	58.425	-8.200
56	ASN C	-6.812	57.894	-8.305	56	ASN D	-5.184	56.866	-7.478
57	PRD H	-6.362	56.261	-9.258	57	PRD CG	-7.123	55.257	-11.177
57	PRD CD	-7.384	56.433	-10.272	57	PRD CB	-6.644	54.170	-10.235
57	PRD CA	-5.679	54.961	-9.332	57	PRD C	-4.381	55.082	-9.946
57	PRD D	-3.509	56.128	-9.945	58	PHE H	-3.998	56.262	-10.491
58	PHE CA	-2.747	56.577	-11.222	58	PHE C	-1.712	57.129	-10.253
58	PHE D	-0.635	57.497	-10.680	58	PHE CS	-2.943	57.582	-12.423
58	PHE CG	-3.983	56.968	-13.357	58	PHE CD1	-3.756	55.788	-14.059
58	PHE CD2	-5.211	57.630	-13.459	58	PHE CE1	-6.722	55.255	-14.928
58	PHE CE2	-6.194	57.895	-14.276	58	PHE CZ	-5.949	55.939	-15.051
59	GLN H	-2.044	57.119	-8.998	59	GLN CA	-1.172	57.583	-7.934
59	GLN C	-0.807	56.403	-7.800	59	GLN D	-1.639	56.883	-6.115
59	GLN CB	-1.062	58.668	-7.889	59	GLN CG	-0.942	59.261	-6.034
59	GLN CD	-3.790	60.157	-5.150	59	GLN DE1	-1.404	61.288	-4.836
59	GLN ME2	-2.959	59.685	-4.742	60	ASP H	0.410	55.895	-7.211
60	ASP CA	0.851	54.792	-6.384	60	ASP C	1.631	55.267	-5.090
60	ASP D	2.827	55.550	-5.231	60	ASP CB	1.596	53.744	-7.188
60	ASP CG	2.077	52.538	-6.380	60	ASP DD1	1.744	52.337	-5.190
60	ASP DD2	2.915	51.841	-7.830	61	ASN H	0.959	55.265	-3.950
61	ASN HD2	-1.364	57.747	-2.347	61	ASN DD1	0.666	58.566	-2.875
61	ASN CG	-8.040	57.670	-2.399	61	ASN CB	0.531	56.401	-1.784
61	ASN CA	1.557	55.734	-2.700	61	ASN C	2.291	54.632	-1.940
61	ASN D	2.933	54.862	-8.902	62	ASN H	2.210	53.434	-2.468
62	ASN CA	2.877	52.368	-1.709	62	ASN C	4.124	51.893	-2.479
62	ASN D	4.951	51.313	-1.770	62	ASN CB	1.783	51.319	-1.421
62	ASN CG	2.371	50.103	-8.497	62	ASN DD1	2.633	49.877	-1.343
62	ASN HD2	2.622	50.208	0.601	63	SER H	4.152	52.184	-3.741
63	SER CA	5.189	51.696	-4.709	63	SER C	5.871	50.256	-5.209
63	SER D	5.533	49.790	-6.249	63	SER CB	6.523	51.958	-4.812
63	SER DG	6.871	50.498	-3.418	64	HIS H	4.202	49.475	-4.639
64	HIS CA	3.994	48.855	-4.935	64	HIS C	3.366	47.759	-6.241
64	HIS D	3.861	46.974	-7.108	64	HIS CB	3.184	47.581	-3.747
64	HIS CG	3.144	46.821	-3.726	64	HIS HD1	2.187	45.247	-4.241
64	HIS CD2	4.854	45.194	-3.135	64	HIS CE1	2.416	43.966	-4.954
64	HIS ME2	3.556	43.920	-3.368	65	GLY H	2.287	48.428	-6.587
65	GLY CA	1.552	48.264	-7.830	65	GLY C	2.392	48.636	-9.037
65	GLY D	2.230	48.078	-10.134	66	TMR H	3.233	49.659	-8.832
66	TMR CA	4.064	58.117	-9.954	66	TMR C	5.889	49.009	-10.291
66	TMR D	5.333	48.789	-11.461	66	TMR CS	4.744	51.511	-9.667
66	TMR DG1	3.637	52.425	-9.406	66	TMR CG2	5.536	52.078	-10.849
67	MIS H	5.685	48.443	-9.274	67	MIS CA	6.783	47.361	-9.458
67	MIS C	6.091	46.141	-10.143	67	MIS D	6.649	65.638	-11.158
67	MIS CB	7.308	47.871	-8.064	67	MIS CG	8.595	46.275	-8.148
67	MIS HD1	8.590	44.907	-8.276	67	MIS CD2	9.904	46.678	-8.076
67	MIS CE1	9.857	44.491	-8.299	67	MIS ME2	10.478	45.514	-8.186
68	VAL H	4.892	45.749	-9.731	68	VAL CA	4.142	44.697	-10.266
68	VAL C	3.856	44.860	-11.740	68	VAL D	4.114	43.942	-12.535
68	VAL CB	2.939	44.252	-9.386	68	VAL CG1	1.988	43.268	-10.820
68	VAL CG2	3.319	43.705	-8.888	69	ALA H	3.373	46.849	-12.113
69	ALA CA	3.837	46.468	-13.429	69	ALA C	4.193	46.398	-14.411
69	ALA D	4.828	45.913	-15.565	69	ALA CS	2.332	47.851	-13.386
70	GLY H	5.348	46.782	-15.914	70	GLY CA	6.595	46.885	-14.670
70	GLY C	7.846	45.370	-15.821	70	GLY D	7.404	45.154	-16.119
71	TMR H	6.820	44.431	-14.138	71	TMR CA	7.177	43.819	-16.446
71	TMR C	6.224	42.586	-15.543	71	TMR D	6.682	41.828	-16.495
71	TMR CB	7.119	42.870	-13.191	71	TMR DG1	8.191	42.592	-12.390

71	YMR CG2	7.274	40.503	-13.596	72	VAL W	4.930	42.007	-15.427
72	VAL CA	3.976	42.491	-16.484	72	VAL C	4.312	43.004	-17.031
72	VAL D	4.341	42.300	-18.860	72	VAL CB	2.516	42.067	-16.085
72	VAL CG1	1.512	42.400	-17.170	72	VAL CG2	2.142	42.327	-14.723
73	ALA W	4.504	44.417	-17.880	73	ALA CA	4.507	45.091	-19.167
73	ALA C	5.433	44.333	-19.355	73	ALA D	5.062	47.100	-20.216
73	ALA CB	3.107	45.441	-19.433	74	ALA W	6.544	48.429	-18.435
74	ALA CA	7.470	47.591	-18.959	74	ALA C	7.740	47.648	-20.342
74	ALA D	7.959	46.640	-21.054	74	ALA CB	8.653	47.446	-17.925
75	LEU W	7.650	48.784	-21.039	75	LEU CA	7.012	48.960	-22.456
75	LEU C	9.192	48.568	-22.966	75	LEU D	10.162	48.750	-22.253
75	LEU CB	7.548	50.471	-22.009	75	LEU CG	6.123	50.913	-22.379
75	LEU CD1	6.079	52.436	-22.300	75	LEU CD2	5.096	50.462	-23.405
76	ASN W	9.147	48.103	-24.169	76	ASN WD2	12.305	46.432	-26.304
76	ASN DD1	10.950	45.840	-27.928	76	ASN CG	11.195	46.274	-26.802
76	ASN CB	10.010	46.651	-25.908	76	ASN CA	10.359	47.738	-24.930
76	ASN C	10.783	49.048	-25.643	76	ASN D	10.157	49.479	-26.619
77	ASN W	11.004	49.664	-25.071	77	ASN CA	12.220	50.957	-25.601
77	ASN C	13.707	51.029	-25.340	77	ASN D	14.364	49.979	-25.313
77	ASN CB	11.335	52.076	-25.117	77	ASN CG	11.250	52.027	-23.616
77	ASN DD1	12.032	51.346	-22.917	77	ASN WD2	10.294	52.741	-23.025
78	SER W	14.125	52.267	-25.164	78	SER CA	15.513	52.614	-24.906
78	SER C	15.810	52.742	-23.436	78	SER D	16.902	53.071	-23.164
78	SER CB	15.905	53.941	-25.517	78	SER CG	15.926	53.870	-26.999
79	ILE W	14.050	52.565	-22.529	79	ILE CA	15.155	52.704	-21.120
79	ILE C	14.617	51.683	-20.230	79	ILE D	13.043	50.041	-20.679
79	ILE CB	14.471	54.174	-20.697	79	ILE CG1	12.945	54.032	-20.814
79	ILE CG2	14.997	55.320	-21.612	79	ILE CD1	12.135	55.176	-20.155
80	GLY W	14.995	51.768	-18.981	80	GLY CA	14.476	50.940	-17.913
80	GLY C	14.612	49.448	-18.219	80	GLY D	15.719	48.994	-18.544
81	VAL W	13.513	48.766	-17.980	81	VAL CA	13.411	47.286	-18.061
81	VAL C	12.511	46.919	-19.217	81	VAL D	12.260	47.739	-20.117
81	VAL CB	13.001	46.755	-16.677	81	VAL CG1	14.030	47.084	-15.573
81	VAL CG2	11.630	47.261	-16.231	82	LEU W	12.126	45.645	-19.216
82	LEU CA	11.312	45.020	-20.256	82	LEU C	10.390	44.028	-19.510
82	LEU D	10.850	43.356	-18.600	82	LEU CB	12.206	44.219	-21.229
82	LEU CG	11.430	43.568	-22.366	82	LEU CD1	10.796	44.657	-23.223
82	LEU CD2	12.359	42.675	-23.192	83	GLY W	9.131	44.100	-19.816
83	GLY CA	8.133	43.321	-19.114	83	GLY C	8.027	42.011	-19.925
83	GLY D	8.946	41.822	-21.026	84	VAL W	7.272	41.112	-19.203
84	VAL CA	6.973	39.807	-19.880	84	VAL C	6.164	40.030	-21.140
84	VAL D	6.424	39.472	-22.194	84	VAL CB	6.256	38.920	-18.041
84	VAL CG1	5.680	37.677	-19.557	84	VAL CG2	7.190	38.507	-17.705
85	ALA W	5.156	40.926	-21.024	85	ALA CA	4.217	41.194	-22.150
85	ALA C	4.213	42.603	-22.396	85	ALA D	3.260	43.401	-22.030
85	ALA CB	2.046	40.663	-21.748	86	PRO W	5.240	43.186	-23.059
86	PRO CA	5.413	44.635	-23.205	86	PRO C	4.321	45.371	-23.947
86	PRO D	4.291	46.605	-23.849	86	PRO CB	6.822	44.784	-23.813
86	PRO CG	7.030	43.466	-24.546	86	PRO CD	6.377	42.440	-23.636
87	SER W	3.548	44.674	-24.769	87	SER CA	2.409	45.324	-25.529
87	SER C	1.103	45.132	-24.897	87	SER D	0.162	45.513	-25.619
87	SER CB	2.401	44.777	-26.927	87	SER CG	3.591	45.143	-27.503
88	ALA W	1.017	44.566	-23.742	88	ALA CB	-0.163	43.510	-21.020
88	ALA CA	-0.273	44.353	-23.004	88	ALA C	-0.090	45.717	-22.690
88	ALA D	-0.174	46.717	-22.435	89	SER W	-2.219	45.691	-22.678
89	SER CG	-4.146	47.102	-24.200	89	SER CG	-4.343	46.903	-22.090
89	SER CA	-3.001	46.867	-22.227	89	SER C	-3.136	46.780	-20.727
89	SER D	-3.793	45.864	-20.209	90	LEU W	-2.446	47.656	-20.037
90	LEU CA	-2.378	47.667	-18.593	90	LEU C	-3.403	48.430	-17.864
90	LEU D	-3.502	49.604	-18.215	90	LEU CB	-0.951	49.273	-18.426
90	LEU CG	-0.233	47.051	-17.174	90	LFU CD1	-0.026	46.361	-17.219
90	LEU CD2	1.160	49.524	-17.047	91	TYR W	-4.266	47.944	-16.938
91	TYR CA	-5.258	48.670	-16.137	91	TYR C	-4.073	48.750	-16.085

-35-

91	TYR D	-4.496	47.749	-14.823	91	TYR CD	-6.686	48.893	-10.314
91	TYR CG	-7.894	48.237	-17.741	91	TYR CD1	-6.595	47.415	-18.755
91	TYR CD2	-7.971	49.275	-18.149	91	TYR CE1	-6.985	47.572	-20.898
91	TYR CE2	-8.315	49.421	-19.492	91	TYR C2	-7.794	48.582	-20.463
91	TYR DM	-8.182	48.752	-21.764	92	ALA M	-4.895	49.958	-14.104
92	ALA CA	-4.549	50.199	-12.787	92	ALA C	-5.823	50.833	-11.903
92	ALA D	-6.723	50.898	-12.050	92	ALA CD	-3.997	51.621	-12.488
93	VAL M	-5.959	48.993	-11.129	93	VAL CA	-7.183	48.854	-18.325
93	VAL C	-6.708	49.814	-8.899	93	VAL D	-6.181	47.993	-8.372
93	VAL CD	-7.957	47.555	-10.611	93	VAL CD1	-9.213	47.488	-9.725
93	VAL CD2	-8.195	47.378	-12.872	94	LVS M	-6.987	50.217	-8.327
94	LVS CA	-6.378	50.464	-6.999	94	LVS C	-7.333	49.985	-5.894
94	LVS D	-8.458	50.480	-5.783	94	LVS CD	-6.051	51.976	-6.818
94	LVS CG	-5.394	52.320	-5.467	94	LVS CD	-4.868	53.785	-5.582
94	LVS CE	-6.399	54.208	-4.199	94	LVS M2	-3.735	55.544	-4.387
95	VAL M	-6.909	49.071	-5.026	95	VAL CA	-7.646	48.457	-3.920
95	VAL C	-6.919	48.699	-2.568	95	VAL D	-7.425	48.156	-1.581
95	VAL CD	-8.104	47.838	-4.319	95	VAL CD1	-8.868	46.852	-5.619
95	VAL CD2	-6.900	46.180	-4.332	96	LEU M	-5.676	48.974	-2.684
96	LEU CA	-4.782	49.183	-1.486	96	LEU C	-4.331	50.559	-1.321
96	LEU D	-3.942	51.121	-2.336	96	LEU CD	-3.589	48.241	-1.573
96	LEU CG	-3.593	46.799	-2.072	96	LEU CD1	-2.287	46.184	-2.163
96	LEU CD2	-4.489	46.082	-1.845	97	GLY M	-4.324	50.975	-0.886
97	GLY CA	-3.890	52.307	0.287	97	GLY C	-2.363	52.437	0.385
97	GLY D	-1.619	51.463	0.165	98	ALA M	-1.954	53.648	0.758
98	ALA CD	-0.428	55.678	1.518	98	ALA CA	-0.563	54.868	0.965
98	ALA C	0.188	53.118	1.917	98	ALA D	1.393	52.921	1.663
99	ASP M	-0.504	52.573	2.912	99	ASP CD2	-2.631	51.842	6.151
99	ASP DD1	-2.730	58.902	4.883	99	ASP CG	-2.883	51.131	5.048
99	ASP CD	-0.648	51.693	5.175	99	ASP CA	0.181	51.610	3.855
99	ASP C	0.166	50.165	3.320	99	ASP D	0.735	49.313	4.829
100	GLY M	-0.424	49.883	2.168	100	GLY CA	-0.343	48.521	1.615
100	GLY C	-1.528	47.651	2.002	100	GLY D	-1.649	46.512	1.479
101	SER M	-2.542	48.128	2.988	101	SER CA	-3.542	47.388	3.315
101	SER C	-4.759	47.894	2.532	101	SER D	-4.758	48.972	1.907
101	SER CD	-3.716	47.447	4.817	101	SER CD	-4.411	48.634	5.289
102	GLY M	-5.821	47.892	2.577	102	GLY CA	-7.877	47.422	1.896
102	GLY C	-8.166	46.536	2.528	102	GLY D	-7.888	45.431	3.030
103	GLN M	-9.377	47.858	2.498	103	GLN CA	-10.535	46.297	3.820
103	GLN C	-10.963	45.232	2.022	103	GLN	-10.779	45.482	0.817
103	GLN CD	-11.671	47.307	3.274	103	GLN CG	-11.368	48.805	4.586
103	GLN CD	-12.360	49.104	4.915	103	GLN DE1	-12.159	49.816	5.902
103	GLN ME2	-13.419	49.197	6.112	104	TYR M	-11.611	44.141	2.451
104	TYR CA	-12.868	43.124	1.588	104	TYR C	-13.031	43.690	0.473
104	TYR D	-12.939	43.276	-0.687	104	TYR CD	-12.697	41.846	2.143
104	TYR CG	-13.629	40.829	2.472	104	TYR CD1	-11.819	39.789	3.377
104	TYR CD2	-10.379	40.959	1.860	104	TYR CE1	-10.809	38.885	3.787
104	TYR CE2	-9.352	48.057	2.171	104	TYR C2	-9.564	39.822	3.881
104	TYR DM	-8.481	38.191	3.324	105	SER M	-13.089	44.572	8.983
105	SER CA	-14.877	45.166	-0.834	105	SER C	-14.172	45.920	-1.159
105	SER D	-14.759	45.935	-2.258	105	SER CD	-15.880	46.121	0.601
105	SER CD	-15.289	47.839	1.450	106	TRP M	-13.879	46.625	-0.834
106	TRP CA	-12.421	47.391	-1.948	106	TRP C	-11.895	46.436	-3.812
106	TRP D	-12.821	46.648	-4.245	106	TRP CD	-11.321	48.254	-1.355
106	TRP CG	-11.845	49.111	-8.206	106	TRP CD1	-12.862	49.524	0.264
106	TRP CD2	-10.658	49.812	0.581	106	TRP ME1	-12.691	50.358	1.360
106	TRP CE2	-11.359	50.573	1.561	106	TRP CE3	-9.275	49.852	0.576
106	TRP C22	-10.671	51.318	2.500	106	TRP C23	-8.568	50.563	1.525
106	TRP CM2	-9.293	51.291	2.455	107	ILE M	-11.339	45.338	-2.681
107	ILE CA	-10.765	44.250	-3.325	107	ILE C	-11.955	43.594	-4.190
107	ILE D	-11.895	43.474	-5.398	107	ILE CD	-9.946	43.183	-2.523
107	ILE CD1	-8.634	43.784	-1.936	107	ILE CD2	-9.632	41.930	-3.381
107	ILE CD1	-8.233	42.998	-8.627	109	ILE M	-12.994	43.292	-3.577

108	IIE CA	-14.114	42.722	-4.321	108	IIE C	-14.439	43.494	-5.386
108	IIE D	-14.894	43.320	-6.552	108	IIE CO	-15.244	42.265	-3.320
108	IIE CG1	-14.726	41.877	-2.482	108	IIE CG2	-16.568	42.824	-4.893
108	IIE CD1	-15.452	48.865	-1.131	109	ASN M	-14.751	46.958	-4.981
109	ASN CA	-15.204	46.010	-5.916	109	ASN C	-14.232	46.867	-7.084
109	ASN B	-14.860	46.272	-8.235	109	ASN CB	-15.280	47.359	-5.207
109	ASN CG	-16.528	47.606	-6.353	109	ASN CD1	-17.655	46.495	-4.646
109	ASN MD2	-16.633	48.447	-3.442	110	GLY M	-12.951	45.908	-6.774
110	GLY CA	-11.952	45.917	-7.865	110	GLY C	-12.108	44.712	-8.812
110	GLY D	-11.929	44.929	-10.834	111	IIE M	-12.379	43.539	-8.246
111	IIE CA	-12.603	42.334	-9.099	111	IIE C	-13.859	42.560	-9.942
111	IIE B	-13.921	42.384	-11.148	111	IIE CB	-12.734	40.948	-8.364
111	IIE CG1	-11.421	40.501	-7.655	111	IIE CG2	-13.122	39.791	-9.347
111	IIE CD1	-11.588	39.786	-6.336	112	GLU M	-14.893	43.875	-9.280
112	GLU CA	-16.118	43.376	-10.846	112	GLU C	-15.872	44.347	-11.171
112	GLU D	-16.467	44.130	-12.246	112	GLU CB	-17.229	43.899	-9.141
112	GLU CG	-17.847	42.917	-8.135	112	GLU CD	-18.724	41.824	-8.685
112	GLU DE1	-19.841	40.866	-8.816	112	GLU DE2	-19.123	41.928	-9.866
113	TRP M	-15.094	45.403	-10.971	113	TRP CA	-16.756	46.408	-12.080
113	TRP C	-14.876	45.663	-13.160	113	TRP D	-14.319	45.932	-14.332
113	TRP CB	-13.882	47.553	-11.434	113	TRP CG	-13.486	48.556	-12.481
113	TRP CD1	-14.148	49.736	-12.681	113	TRP CD2	-12.441	48.552	-13.463
113	TRP ME1	-13.597	50.443	-13.723	113	TRP CE2	-12.545	49.761	-14.215
113	TRP CE3	-11.451	47.645	-13.809	113	TRP CZ2	-11.696	50.045	-15.274
113	TRP CZ3	-10.610	47.899	-14.879	113	TRP CH2	-10.752	49.074	-15.603
114	ALA M	-13.089	44.801	-12.832	114	ALA CA	-12.333	44.865	-13.874
114	ALA C	-13.199	43.179	-14.752	114	ALA D	-12.963	43.874	-15.978
114	ALA CB	-11.299	43.192	-13.140	115	IIE M	-14.174	42.540	-14.119
115	IIE CA	-13.870	41.640	-14.897	115	IIE C	-15.928	42.485	-15.856
115	IIE D	-14.877	42.225	-17.870	115	IIE CB	-16.880	40.840	-13.922
115	IIE CG1	-15.218	39.836	-13.843	115	IIE CG2	-17.151	40.168	-14.755
115	IIE CD1	-16.004	39.411	-11.743	116	ALA M	-16.534	43.527	-15.267
116	ALA CA	-17.390	44.440	-16.850	116	ALA C	-16.786	45.049	-17.278
116	ALA D	-17.323	45.255	-18.343	116	ALA CB	-18.011	45.510	-15.151
117	ASN M	-15.423	45.390	-17.122	117	ASN CA	-16.553	45.967	-18.139
117	ASN C	-13.827	46.974	-19.834	117	ASN D	-12.997	45.436	-19.820
117	ASN CB	-13.615	46.958	-17.426	117	ASN CG	-14.408	48.177	-16.939
117	ASN OD1	-14.565	47.882	-17.773	117	ASN MD2	-14.931	48.249	-15.736
118	ASN M	-14.223	43.725	-18.967	118	ASN CA	-13.760	42.642	-19.832
118	ASN C	-12.240	42.444	-19.843	118	ASN D	-11.617	42.309	-20.932
118	ASN CB	-14.247	42.863	-21.279	118	ASN CG	-15.737	43.060	-21.395
118	ASN OD1	-16.510	42.321	-20.759	118	ASN MD2	-16.136	44.096	-22.133
119	MET M	-11.686	42.500	-18.675	119	MET CA	-10.232	42.222	-18.478
119	MET C	-10.025	40.734	-18.928	119	MET D	-10.888	39.838	-18.759
119	MET CB	-9.810	42.461	-17.055	119	MET CG	-9.880	43.883	-16.582
119	MET SD	-8.788	44.943	-17.526	119	MET CE	-9.982	46.861	-18.263
120	ASP M	-8.904	48.437	-19.584	120	ASP CA	-8.480	39.118	-20.830
120	ASP C	-7.822	38.390	-18.854	120	ASP D	-8.838	37.189	-18.490
120	ASP CB	-7.555	39.156	-21.236	120	ASP CG	-8.237	39.730	-22.454
120	ASP OD1	-7.881	40.706	-23.884	120	ASP DD2	-9.327	39.135	-22.739
121	VAL M	-7.071	39.117	-18.115	121	VAL CA	-6.224	38.601	-16.974
121	VAL C	-6.296	39.534	-15.786	121	VAL D	-6.284	40.788	-15.909
121	VAL CB	-4.755	38.587	-17.496	121	VAL CG1	-3.758	38.174	-16.627
121	VAL CG2	-4.787	37.916	-18.846	122	IIE M	-6.318	38.978	-14.590
122	IIE CA	-6.248	39.799	-13.397	122	IIE C	-5.828	39.262	-12.627
122	IIE D	-4.829	38.812	-12.469	122	IIE CB	-7.476	39.604	-12.466
122	IIE CG1	-8.686	40.392	-13.063	122	IIE CG2	-7.221	39.883	-10.954
122	IIE CD1	-9.976	39.788	-12.393	123	ASN M	-6.263	40.222	-12.110
123	ASN CA	-3.145	39.854	-11.232	123	ASN C	-3.582	48.404	-9.861
123	ASN B	-3.708	41.631	-9.833	123	ASN CB	-1.828	40.478	-11.697
123	ASN CG	-8.692	40.848	-10.777	123	ASN CD1	-8.863	38.990	-11.818
123	ASN MD2	-8.346	40.747	-9.728	124	MET M	-3.458	39.604	-8.832
124	MET CA	-3.658	39.973	-7.438	124	MET C	-2.423	39.603	-8.614

124	NET D	-2.304	30.508	-6.093	124	NET CB	-4.943	39.387	-4.010
124	NET CG	-6.198	40.082	-7.473	124	NET SC	-7.181	39.472	-6.450
124	NET CL	-7.949	30.095	-7.842	125	SEN M	-1.454	40.496	-6.502
125	SEN CA	-0.193	40.287	-5.749	125	SEN C	-0.412	40.712	-6.326
125	SEN D	0.239	41.617	-3.055	125	SEN CB	1.021	41.027	-6.328
125	SEN DC	1.444	40.496	-7.575	126	LEU M	-1.433	40.073	-3.778
126	LEU CA	-1.842	40.347	-2.386	126	LEU C	-2.438	39.056	-1.807
126	LEU D	-2.964	38.136	-2.529	126	LEU CB	-2.791	41.568	-2.410
126	LEU CG	-3.988	41.447	-3.333	126	LEU CD1	-3.278	41.131	-2.370
126	LEU CD2	-6.179	42.760	-4.073	127	GLY M	-2.522	39.082	-0.481
127	GLY CA	-3.035	37.071	0.193	127	GLY C	-3.176	38.180	1.632
127	GLY D	-2.446	39.030	2.220	128	GLY M	-4.121	37.443	2.222
128	GLY CA	-4.675	37.496	3.642	128	GLY C	-4.644	36.039	4.104
128	GLY D	-4.993	36.158	3.276	129	PRO M	-4.919	35.057	5.402
129	PRO CA	-4.671	34.523	5.998	129	PRO C	-6.116	34.006	6.082
129	PRO D	-6.338	32.887	6.305	129	PRO CB	-6.040	34.684	7.384
129	PRO CG	-4.419	36.116	7.727	129	PRO CD	-6.239	36.870	6.418
130	SEN M	-7.051	35.019	5.912	130	SEN CA	-0.670	34.611	6.023
130	SEN C	-9.218	34.084	4.726	130	SEN D	-0.949	35.001	4.020
130	SEN CB	-9.049	35.351	7.216	130	SEN DC	-0.723	34.624	8.403
131	GLY M	-10.083	33.967	4.349	131	GLY CA	-10.824	34.229	3.874
131	GLY C	-12.205	34.713	3.562	131	GLY D	-12.495	34.722	4.751
132	SEN M	-13.940	33.058	2.594	132	SEN CA	-14.607	33.433	3.011
132	SEN C	-15.289	34.905	1.936	132	SEN D	-14.709	34.986	0.824
132	SEN CB	-14.890	36.927	3.145	132	SEN DC	-14.693	37.539	1.875
133	ALA M	-16.547	34.388	2.294	133	ALA CA	-17.507	34.057	1.324
133	ALA C	-17.650	34.945	0.097	133	ALA D	-17.743	34.437	-1.016
133	ALA CB	-18.866	33.828	1.994	133	ALA D	-17.683	36.288	0.294
134	ALA CA	-17.872	37.259	-0.792	134	ALA M	-16.435	37.349	-1.674
134	ALA D	-16.781	37.585	-2.069	134	ALA C	-16.263	39.600	-0.187
135	LEU M	-15.478	37.229	-1.046	135	LEU CA	-14.197	37.244	-1.804
135	LEU C	-14.158	36.005	-2.705	135	LEU D	-13.794	36.020	-3.890
135	LEU CB	-13.038	37.328	-0.798	135	LEU CG	-11.693	37.130	-1.908
135	LEU CD1	-11.460	38.415	-2.292	135	LEU CD2	-10.982	34.807	-0.519
136	LVS M	-14.809	34.823	-2.173	136	LVS CA	-14.563	33.597	-3.013
136	LVS C	-15.544	33.739	-4.150	136	LVS C	-19.279	33.431	-5.305
136	LVS CB	-14.903	32.341	-2.186	136	LVS CG	-14.743	31.067	-3.043
136	LVS CD	-15.083	29.892	-2.134	136	LVS CE	-15.743	28.707	-2.778
136	LVS NZ	-15.308	28.411	-4.160	137	ALA M	-16.744	34.260	-3.047
137	ALA CA	-17.795	34.416	-6.083	137	ALA C	-17.338	35.303	-6.043
137	ALA D	-17.709	35.049	-7.208	137	ALA CB	-19.094	34.941	-6.263
138	ALA M	-16.529	36.301	-5.729	138	ALA D	-16.001	37.311	-0.685
138	ALA C	-14.903	36.696	-7.937	138	ALA CB	-14.985	36.043	-8.762
138	ALA CB	-15.522	38.967	-5.034	139	VAL M	-13.950	33.959	-7.027
139	VAL CA	-12.946	35.291	-7.037	139	VAL C	-13.423	34.228	-0.720
139	VAL D	-13.208	34.070	-9.877	139	VAL CB	-11.830	34.671	-0.968
139	VAL CG1	-10.919	33.056	-7.066	139	VAL CD2	-11.078	33.780	-0.253
140	ASP M	-14.595	33.336	-8.122	139	VAL CE	-15.274	32.496	-0.929
140	ASP C	-16.023	33.131	-10.084	140	ASP CA	-16.980	32.579	-11.100
140	ASP CB	-16.149	31.549	-1.158	140	ASP D	-15.388	30.648	-7.184
140	ASP CD1	-14.178	30.403	-7.202	140	ASP CG	-15.139	30.132	-6.329
141	LVS M	-14.658	34.263	-9.020	141	LVS CA	-17.373	35.006	-10.868
141	LVS C	-18.373	35.419	-11.946	141	LVS D	-16.700	35.248	-13.111
141	LVS CB	-18.039	36.275	-10.325	141	LVS CG	-18.004	37.096	-11.306
141	LVS CD	-19.596	38.187	-10.536	141	LVS CE	-20.572	39.051	-11.250
141	LVS NZ	-21.138	40.037	-10.275	142	ALA M	-15.167	35.040	-11.566
142	ALA CA	-14.173	36.192	-12.614	142	ALA C	-13.818	35.010	-13.071
142	ALA D	-13.770	35.169	-14.755	142	ALA CB	-12.070	36.097	-11.948
143	VAL M	-13.882	33.806	-12.832	143	VAL CA	-13.160	32.703	-13.650
143	VAL C	-14.346	32.233	-14.496	143	VAL D	-14.140	31.006	-15.639
143	VAL CB	-12.581	31.073	-12.714	143	VAL CD	-12.300	30.370	-13.661
143	VAL CG2	-11.395	32.195	-12.014	143	VAL CE1	-13.531	32.258	-13.075
144	ALA CA	-16.744	31.034	-14.641	144	ALA M	-16.928	32.681	-15.061
144	ALA C				144	ALA C			

144	ALA C	-17.380	32.263	-16.933	144	ALA CB	-17.942	32.968	-13.788
145	SLY M	-16.307	33.948	-15.704	145	SLY CB	-16.682	34.917	-16.786
145	SLY C	-15.609	34.773	-17.829	145	SLY D	-15.910	35.321	-18.873
145	SLY CB	-17.016	36.376	-16.614	145	SLY DC	-15.982	36.935	-18.849
146	GLY M	-16.577	33.086	-17.565	146	GLY CA	-15.619	33.799	-18.678
146	GLY C	-12.273	34.491	-18.395	146	GLY D	-11.420	34.386	-19.266
147	VAL M	-12.150	35.162	-17.254	147	VAL CA	-10.874	35.856	-19.912
147	VAL C	-9.850	34.836	-16.323	147	VAL D	-10.171	35.991	-19.486
147	VAL CB	-11.152	36.977	-15.889	147	VAL CC1	-9.896	37.803	-18.878
147	VAL CC2	-12.340	37.933	-14.230	148	VAL M	-8.983	35.018	-16.683
148	VAL CA	-7.482	34.230	-14.008	148	VAL C	-7.137	34.907	-14.701
148	VAL D	-6.842	36.133	-14.750	148	VAL CB	-6.273	34.126	-14.938
148	VAL CC1	-5.079	33.683	-14.281	148	VAL CC2	-6.590	35.432	-12.249
149	VAL M	-7.258	34.355	-13.531	149	VAL CA	-6.987	34.965	-12.349
149	VAL C	-8.708	34.385	-11.613	149	VAL D	-5.624	35.173	-11.439
149	VAL CB	-8.224	34.890	-12.315	149	VAL CC1	-7.893	35.619	-10.009
149	VAL CC2	-9.456	35.356	-12.096	150	VAL M	-4.732	35.301	-11.404
150	VAL CA	-3.393	34.987	-10.901	150	VAL C	-3.157	35.623	-9.599
150	VAL D	-3.882	36.778	-9.400	150	VAL CB	-2.274	35.369	-11.951
150	VAL CC1	-8.973	34.633	-11.661	150	VAL CC2	-2.675	34.943	-13.301
151	ALA M	-2.968	34.946	-8.595	151	ALA CA	-2.361	35.582	-7.287
151	ALA C	-1.880	35.036	-6.457	151	ALA D	-0.618	35.889	-4.984
151	ALA CB	-3.557	35.398	-6.307	152	ALA M	-0.490	35.987	-5.022
152	ALA CA	-0.714	35.438	-5.212	152	ALA C	0.304	36.320	-4.198
152	ALA D	-0.728	34.466	-3.467	152	ALA CB	1.266	36.607	-4.204
153	ALA M	1.125	33.302	-3.912	153	ALA CA	0.840	32.258	-2.963
153	ALA C	0.931	32.725	-3.511	153	ALA D	0.317	32.192	-0.599
153	ALA CB	1.750	31.038	-3.193	154	GLY M	1.827	33.693	-1.244
154	GLY CA	2.843	34.211	0.123	154	GLY C	3.519	34.669	0.550
154	GLY D	4.189	33.267	-0.118	155	ASH M	3.958	34.788	1.568
155	ASH CA	5.344	34.787	2.037	155	ASH C	5.399	34.258	3.662
155	ASH D	6.101	34.829	4.295	155	ASH CB	6.008	34.198	1.984
155	ASH CC	5.890	36.792	0.500	155	ASH CC1	6.123	34.965	-0.934
155	ASH CC2	5.454	37.965	0.392	156	GLU M	4.711	33.168	3.675
156	GLU CA	4.833	32.537	4.970	156	GLU C	5.522	31.378	5.183
156	GLU D	5.374	30.637	6.222	156	GLU CB	3.203	31.980	5.100
156	GLU CC	2.493	32.642	6.368	156	GLU CD	3.994	33.051	6.278
156	GLU CC1	1.744	34.322	5.312	156	GLU CC2	3.106	34.456	7.146
157	GLY M	6.389	31.057	4.227	157	GLY CA	7.386	29.917	4.387
157	GLY C	6.503	28.622	4.553	157	GLY D	5.416	29.346	4.889
158	TMR M	7.147	27.793	5.382	158	TMR CB	8.079	29.396	3.850
158	TMR CC1	8.707	25.487	6.217	158	TMR C	7.564	29.366	5.296
158	TMR CA	6.952	26.687	5.702	159	SLY M	6.190	26.480	7.187
158	TMR D	6.479	27.335	7.977	159	SLY C	5.338	25.441	7.497
159	SLY DC	3.141	25.904	10.523	159	SLY CB	3.673	26.185	9.212
159	SLY CA	4.831	25.210	8.555	159	SLY C	4.494	23.720	8.946
159	SLY D	3.339	23.281	9.030	160	GLY M	5.574	22.967	8.838
160	GLY CA	5.436	21.504	8.895	160	GLY C	6.576	21.045	7.738
160	GLY D	4.809	21.326	6.555	161	SLY M	3.925	20.318	8.116
161	SLY CA	2.654	19.777	7.054	161	SLY C	1.477	20.788	6.786
161	SLY D	0.656	20.347	5.869	161	SLY CB	2.344	20.293	7.271
161	SLY DC	1.834	18.028	8.595	162	SLY M	1.303	21.841	7.699
162	SLY CA	0.167	22.725	7.113	162	SLY C	0.430	23.552	5.848
162	SLY D	1.533	23.040	5.394	162	SLY CB	-0.213	23.466	8.242
162	SLY DC	0.204	23.891	9.480	163	SLY M	-0.679	23.921	8.197
163	SLY CA	-0.611	24.750	3.990	163	SLY C	-0.441	26.177	4.513
163	SLY D	-1.078	26.548	5.504	163	SLY CB	-1.890	24.642	3.211
163	SLY DC	-1.092	25.718	2.331	164	TMR M	0.387	26.932	3.852
164	TMR CA	0.609	28.340	4.312	164	TMR C	0.105	29.286	3.194
164	TMR D	5.485	29.502	3.278	164	TMR CB	2.095	28.518	4.818
164	TMR CC1	2.984	28.382	3.692	164	TMR CC2	2.397	27.610	6.001
165	VAL M	-0.513	28.742	2.190	165	VAL CA	-0.959	29.562	1.018
165	VAL C	-2.026	28.545	1.497	165	VAL D	-2.929	30.132	1.280

165 VAL CB	-1.339	28.624	-0.161	165 VAL CC1	-1.947	29.397	-1.174
165 VAL CC2	-2.216	27.716	-0.991	166 GLY M	-1.910	31.821	1.129
166 GLY CA	-2.943	32.778	1.626	166 GLY C	-4.098	32.898	0.617
166 GLY D	-6.124	32.394	-0.394	167 TYR M	-0.064	33.739	0.970
167 TYR CA	-0.223	34.046	0.113	167 TYR C	-3.993	33.289	-0.686
167 TYR D	-5.674	36.293	0.884	167 TYR CB	-7.464	34.232	0.964
167 TYR CG	-7.791	32.964	1.709	167 TYR CD1	-7.298	32.783	2.947
167 TYR CD2	-0.710	32.116	1.133	167 TYR CE1	-7.567	31.928	3.619
167 TYR CE2	-9.068	30.933	1.809	167 TYR C2	-0.486	39.671	3.046
167 TYR DM	-4.882	29.681	3.658	168 PRO M	-4.380	33.499	-1.830
168 PRO CG	-6.963	36.376	-3.938	168 PRO CD	-6.273	36.752	-2.624
168 PRO CB	-7.964	33.344	-3.903	168 PRO CA	-7.334	34.637	-2.560
168 PRO C	-6.398	33.336	-3.270	168 PRO D	-7.097	32.520	-3.912
169 GLY M	-5.886	33.193	-3.189	169 GLY CA	-4.446	32.877	-3.927
169 GLY C	-4.937	30.702	-3.470	169 GLY D	-4.880	29.733	-4.249
170 LYS M	-3.402	30.579	-2.293	170 LYS CA	-3.856	29.268	-1.748
170 LYS C	-7.033	28.773	-2.316	170 LYS D	-7.308	27.894	-2.524
170 LYS CB	-6.246	29.294	-0.226	170 LYS CG	-9.795	28.106	0.983
170 LYS CD	-6.230	28.289	2.031	170 LYS CE	-3.731	27.271	3.029
170 LYS M2	-4.239	27.663	3.219	171 TYR M	-7.830	29.616	-3.168
171 TYR CA	-9.012	29.043	-3.939	171 TYR C	-0.683	29.309	-0.113
171 TYR D	-7.740	28.714	-3.928	171 TYR CB	-9.962	30.224	-4.242
171 TYR CG	-10.497	30.984	-3.047	171 TYR CD1	-11.960	30.303	-1.982
171 TYR CD2	-10.436	32.374	-3.026	171 TYR CE1	-11.820	31.003	-0.867
171 TYR CE2	-10.943	33.088	-1.936	171 TYR C2	-11.828	32.396	-0.886
171 TYR DM	-12.808	33.319	0.170	172 PRO M	-9.297	27.204	-3.374
172 PRO CA	-0.093	26.617	-6.396	172 PRO C	-9.233	27.196	-7.089
172 PRO D	-0.323	26.784	-0.881	172 PRO CB	-10.167	28.329	-6.513
172 PRO CG	-10.400	25.271	-0.096	172 PRO CD	-10.364	26.669	-4.516
173 SER M	-10.897	28.167	-0.019	173 SER CA	-10.220	28.818	-9.330
173 SER C	-9.025	29.773	-0.393	173 SER D	-0.966	30.233	-10.742
173 SER CB	-11.528	29.623	-9.481	173 SER CG	-11.993	30.166	-0.406
174 VAL M	-0.162	29.944	-0.614	174 VAL CA	-7.093	30.891	-0.853
174 VAL C	-0.754	30.231	-0.068	174 VAL D	-3.612	29.192	-0.344
174 VAL CB	-6.899	31.775	-7.596	174 VAL CC1	-3.796	32.837	-7.617
174 VAL CC2	-0.220	32.503	-7.323	175 ILE M	-4.913	30.729	-9.883
175 ILE CA	-3.869	30.196	-10.024	175 ILE C	-2.714	30.736	-0.694
175 ILE D	-2.450	31.938	-0.953	175 ILE CB	-2.953	30.524	-11.419
175 ILE CC1	-0.837	29.978	-12.524	175 ILE CC2	-1.491	30.089	-11.312
175 ILE CC1	-3.692	30.529	-13.946	176 ALA M	-2.220	30.028	-7.923
176 ALA CA	-1.335	30.517	-6.870	176 ALA C	0.120	29.301	-7.310
176 ALA D	0.433	29.213	-7.839	176 ALA CB	-1.639	29.838	-0.541
177 VAL M	0.864	31.410	-7.180	177 VAL CA	2.261	31.834	-7.636
177 VAL C	3.223	31.693	-6.473	177 VAL D	3.178	32.657	-3.721
177 VAL CB	2.439	32.607	-0.753	177 VAL CC1	3.842	32.667	-9.392
177 VAL CC2	3.374	32.532	-0.843	178 GLY M	4.077	30.654	-6.358
178 GLY CA	3.168	30.703	-3.339	178 GLY C	6.446	31.233	-6.974
178 GLY D	6.499	31.635	-7.296	179 ALA M	7.512	31.447	-3.287
179 ALA CA	0.719	32.037	-0.359	179 ALA C	9.939	31.099	-5.779
179 ALA D	10.198	30.481	-4.710	179 ALA CB	9.023	33.231	-4.973
180 VAL M	10.639	31.162	-6.883	180 VAL CA	11.970	30.482	-6.981
180 VAL C	13.048	31.885	-7.171	180 VAL D	12.712	32.691	-7.627
180 VAL CB	12.075	29.514	-0.166	180 VAL CC1	11.271	28.231	-7.833
180 VAL CC2	11.675	30.129	-9.300	181 ASP M	14.267	31.293	-4.080
181 ASP CA	19.431	32.188	-7.039	181 ASP C	15.942	31.804	-8.462
181 ASP D	19.339	31.090	-9.292	181 ASP CB	16.466	31.921	-5.916
181 ASP CG	17.120	30.534	-5.971	181 ASP CD1	17.103	29.799	-6.972
181 ASP DD2	17.685	30.284	-4.807	182 SER M	17.087	32.386	-0.847
182 SER CA	17.622	32.214	-10.191	182 SER C	18.193	30.817	-10.496
182 SER D	18.365	30.692	-11.670	182 SER CB	18.678	33.313	-10.466
182 SER CG	18.016	34.961	-10.473	183 SER M	18.258	30.042	-0.423
183 SER-CA	18.716	28.645	-9.464	183 SER C	17.881	27.614	-9.947
183 SER D	17.839	26.613	-0.397	183 SER CB	19.296	28.323	-0.897



183	SEN GC	29.989	29.813	-0.291	194	ASN W	16.973	28.094	-0.692
184	ASN CA	19.164	27.317	-0.990	194	ASN C	16.991	26.720	-0.197
184	ASN D	14.138	25.789	-0.897	194	ASN CD	19.914	26.941	-10.722
184	ASN CG	14.993	26.998	-12.076	194	ASN CD1	14.788	28.184	-12.277
184	ASN ND2	19.352	24.210	-13.076	193	GLN W	19.942	27.247	-7.199
183	GLN CA	19.276	24.846	-9.939	193	GLN C	14.200	27.694	-5.293
183	GLN D	14.159	28.726	-9.386	193	GLN CD	16.599	26.968	-5.191
183	GLN CG	16.339	26.242	-9.614	193	GLN CD	18.011	26.182	-9.206
183	GLN DE1	18.864	29.799	-6.961	193	GLN NE2	11.266	26.386	-1.934
186	ARG W	13.278	26.938	-6.648	186	ARG CA	12.185	27.774	-3.841
186	ARG C	12.780	28.782	-2.866	186	ARG D	13.698	28.384	-2.893
186	ARG CD	11.215	26.043	-9.116	186	ARG CG	10.214	27.471	-2.161
186	ARG CD	9.467	26.337	-1.668	186	ARG NE	9.866	26.333	-0.117
186	ARG CZ	9.961	26.879	1.039	186	ARG NM1	9.367	27.880	1.698
186	ARG NM2	10.966	26.721	1.783	197	ALA W	12.294	30.889	-2.833
187	ALA CA	12.728	31.064	-1.095	197	ALA C	12.262	30.604	-0.317
187	ALA D	11.158	30.043	-0.387	197	ALA CD	12.144	32.402	-2.344
188	SEN W	13.891	30.770	0.849	198	SEN CA	12.671	30.286	1.868
188	SEN C	11.396	30.847	2.412	198	SEN D	10.740	30.111	3.212
188	SEN CD	13.767	30.456	2.937	198	SEN CG	14.137	31.826	2.841
189	PME W	10.943	32.010	1.974	199	PME CA	9.697	32.688	2.418
189	PME C	8.499	32.198	1.609	199	PME D	7.389	32.556	2.011
189	PME CD	9.787	34.217	2.243	199	PME CG	10.117	34.696	0.867
189	PME CD1	9.147	34.930	-0.121	199	PME CD2	11.418	35.116	0.667
189	PME CD1	9.483	33.187	-1.611	199	PME CD2	11.769	35.545	-0.781
189	PME C2	10.786	35.886	-1.729	199	SEN W	8.703	31.926	0.499
190	SEN CA	7.626	31.896	-0.391	199	SEN C	6.663	30.162	0.328
190	SEN D	7.934	29.023	0.866	199	SEN CD	8.181	30.590	-1.788
190	SEN CG	7.136	30.337	-2.618	191	SEN W	6.388	30.931	0.224
191	SEN CA	4.341	29.676	0.987	191	SEN C	4.261	28.330	0.223
191	SEN D	4.543	28.269	-0.895	191	SEN CD	3.015	30.611	0.911
191	SEN CG	2.729	31.289	1.934	192	VAL W	3.786	27.310	0.928
192	VAL CA	3.629	25.932	0.391	192	VAL C	2.294	25.291	0.686
192	VAL D	1.999	25.698	1.998	192	VAL CD	4.781	25.127	1.888
192	VAL CG1	6.144	25.727	0.722	192	VAL CG2	4.617	25.104	2.992
193	GLY W	3.938	24.172	0.047	193	GLY CA	0.629	23.944	0.410
193	GLY C	0.081	23.029	-0.901	193	GLY D	0.530	23.244	-2.819
194	PRC W	-1.023	22.289	-0.721	194	PRC CA	-1.662	21.651	-1.873
194	PRC C	-2.237	22.609	-2.914	194	PRC D	-2.403	22.244	-4.085
194	PRC CD	-2.769	20.783	-1.210	194	PRC CG	-2.311	20.622	0.213
194	PRC CD	-1.633	21.954	0.578	195	GLU W	-2.822	23.793	-2.439
195	GLU CA	-3.145	24.890	-3.282	195	GLU C	-2.093	23.631	-4.058
195	GLU D	-2.316	26.298	-4.936	195	GLU CD	-4.043	23.786	-2.670
195	GLU CG	-4.942	25.124	-1.635	195	GLU CD	-6.318	24.860	-0.180
195	GLU DE1	-3.110	24.960	0.163	195	GLU DE2	-5.138	24.520	0.785
196	LEU W	-0.529	25.264	-3.070	196	LEU CA	0.241	25.929	-4.664
196	LEU C	0.228	25.376	-6.059	196	LEU D	0.303	24.121	-6.183
196	LEU CD	1.340	25.739	-3.894	196	LEU CG	2.770	26.178	-4.643
196	LEU CD1	2.739	27.716	-4.638	196	LEU CD2	4.027	25.721	-3.911
197	ASP W	0.140	26.208	-7.093	197	ASP CA	0.832	25.774	-8.480
197	ASP C	1.307	25.738	-9.293	197	ASP D	1.053	24.734	-9.914
197	ASP CD	-1.067	26.598	-9.191	197	ASP CG	-2.486	26.351	-8.549
197	ASP CD1	-2.804	28.155	-8.934	197	ASP CD2	-3.033	27.327	-8.088
198	VAL W	2.813	26.889	-9.344	198	VAL CA	3.206	26.978	-10.209
198	VAL C	4.187	27.930	-9.514	198	VAL D	3.752	28.699	-8.587
198	VAL CD	2.594	27.476	-11.627	198	VAL CG1	1.938	26.726	-12.937
198	VAL CG2	1.337	28.919	-11.484	199	MEY W	8.374	27.916	-18.016
199	MEY CA	6.439	28.802	-9.498	199	MEY C	6.843	29.810	-18.578
199	MEY D	6.696	29.318	-11.793	199	MEY CD	7.660	27.978	-9.977
199	MEY CG	7.363	26.849	-8.139	199	MEY CD	6.753	27.449	-6.968
199	MEY CD	8.227	27.751	-8.887	200	ALA W	7.426	30.942	-18.183
200	ALA CA	7.991	31.929	-11.035	200	ALA C	9.888	32.606	-18.272
200	ALA D	9.127	32.924	-0.860	200	ALA CD	6.932	32.078	-11.638

201	PRC B	9.927	33.499	-18.991	201	PRC CA	11.813	36.130	-18.238
201	PRD C	10.430	35.127	-9.233	201	PRC D	9.579	35.987	-9.682
201	PRD CB	11.837	36.723	-11.409	201	PRC CG	11.392	36.960	-12.678
201	PRD CD	9.841	33.616	-12.493	202	GLY B	10.929	35.294	-8.821
202	GLY CA	10.473	36.254	-7.844	202	GLY C	11.589	36.638	-6.115
202	GLY D	11.332	37.124	-4.979	203	VAL B	12.813	36.303	-6.613
203	VAL CA	13.048	36.929	-3.716	203	VAL C	14.796	39.017	-6.469
203	VAL C	13.133	37.731	-7.593	203	VAL CD	14.814	35.688	-3.351
203	VAL CG1	14.896	36.106	-4.632	203	VAL CG2	14.879	36.741	-4.378
204	SEP B	14.665	39.182	-3.859	204	SEP CA	15.572	40.281	-6.487
204	SEP C	15.067	40.619	-7.872	204	SEP C	15.786	40.693	-2.889
204	SEP CB	17.987	39.976	-6.326	204	SEP CD	17.732	41.186	-6.472
205	ILE B	13.771	40.863	-8.008	205	ILE CA	13.869	41.234	-9.223
205	ILE C	13.207	42.749	-9.478	205	ILE D	12.676	43.498	-3.648
205	ILE CB	11.832	40.833	-9.144	205	ILE CG1	11.436	39.336	-8.810
205	ILE CG2	10.899	41.281	-10.467	205	ILE CD1	12.257	38.412	-9.771
206	GLN B	13.934	43.093	-10.489	206	GLN CA	14.204	44.517	-10.234
206	GLN C	13.002	44.978	-11.630	206	GLN D	12.669	44.318	-12.621
206	GLN CB	13.453	44.703	-11.740	206	GLN CG	16.684	44.163	-10.980
206	GLN CD	17.283	45.145	-10.007	206	GLN DE1	18.328	44.934	-9.353
206	GLN DE2	14.534	46.260	-9.857	207	SEP B	12.359	46.864	-11.214
207	SEP CA	11.217	46.571	-11.927	207	SEP C	11.089	48.093	-11.749
207	SEP D	11.919	48.437	-11.004	207	SEP CB	9.918	43.833	-11.569
207	SEP CG	8.993	46.034	-12.613	208	TMR B	10.894	48.664	-12.326
208	TMR CG2	9.171	50.339	-14.734	208	TMR CD1	7.570	49.414	-13.144
208	TMR CB	8.620	50.413	-13.357	208	TMR CA	9.673	50.092	-12.173
208	TMR C	9.197	50.488	-10.803	208	TMR D	8.423	49.807	-10.849
209	LEU B	9.634	51.613	-10.228	209	LEU CA	9.192	52.158	-8.939
209	LEU C	8.673	53.610	-9.262	209	LEU D	9.140	54.227	-19.222
209	LEU CB	10.333	52.192	-7.938	209	LEU CG	10.804	58.816	-7.616
209	LEU CD1	13.968	51.114	-6.472	209	LEU CD2	9.607	50.282	-6.649
210	PRD B	7.790	54.139	-8.444	210	PRD CA	7.273	55.517	-8.649
210	PRD C	8.383	56.573	-8.439	210	PRC D	9.491	56.445	-8.104
210	PRD CB	6.302	59.733	-7.317	210	PRC CG	6.004	54.379	-6.944
210	PRD CD	7.193	53.491	-7.271	211	SLY B	8.077	57.663	-9.333
211	GLY CA	9.049	58.763	-9.410	211	SLY C	10.094	58.434	-10.690
211	GLY D	11.176	59.003	-10.259	212	ASN B	9.851	57.770	-11.587
212	ASN CA	10.903	57.422	-12.643	212	ASN C	12.039	56.793	-12.056
212	ASN D	13.188	57.181	-12.420	212	ASN CB	11.224	58.399	-12.499
212	ASN CG	11.803	58.183	-14.814	212	ASN CD1	11.853	57.884	-15.323
212	ASN ND2	12.273	59.159	-15.576	213	LVS B	11.803	55.749	-11.247
213	LVS CA	12.810	54.946	-10.337	213	LVS C	12.668	53.439	-10.866
213	LVS D	11.773	53.039	-11.613	213	LVS CB	12.769	55.241	-9.839
213	LVS CG	13.206	54.694	-8.767	213	LVS CD	13.246	57.030	-7.312
213	LVS CE	14.193	59.218	-6.870	213	LVS DE	15.048	58.785	-7.921
214	TTR B	13.681	52.703	-10.444	214	TTR CA	13.800	51.246	-10.722
214	TTR C	14.383	50.600	-9.489	214	TTR D	19.211	51.293	-8.817
214	TTR CB	14.641	50.981	-11.984	214	TTR CG	14.130	51.621	-13.746
214	TTR CD1	14.689	52.047	-13.678	214	TTR CD2	19.129	51.063	-14.814
214	TTR CE1	14.230	53.478	-14.814	214	TTR CE2	12.634	51.669	-15.178
214	TTR CE	13.204	52.893	-15.530	214	TTR DM	12.756	53.458	-16.696
215	GLY B	14.938	49.947	-9.158	215	GLY CA	14.622	48.772	-7.903
215	GLY C	14.130	47.323	-7.749	215	GLY D	19.249	46.917	-8.321
216	ALA B	14.810	46.638	-6.831	216	ALA CA	14.434	45.203	-6.781
216	ALA C	13.882	44.922	-5.912	216	ALA D	13.948	45.527	-4.479
216	ALA CB	15.713	44.354	-6.887	217	TTR B	12.758	43.982	-9.373
217	TTR CA	11.964	43.688	-6.440	217	TTR C	12.833	41.928	-4.547
217	TTR D	12.202	41.442	-5.656	217	TTR CE	18.473	43.862	-6.370
217	TTR CG	10.117	43.293	-4.214	217	TTR CD1	18.846	45.991	-3.236
217	TTR CD2	9.816	43.933	-4.783	217	TTR CE1	18.499	47.267	-2.790
217	TTR CE2	8.634	47.219	-4.381	217	TTR CE2	9.358	47.882	-3.391
217	TTR DM	8.953	49.160	-2.988	218	ASN B	12.790	61.396	-3.391
218	ASN CA	11.640	39.942	-3.227	218	ASN C	10.284	39.636	-2.749

218	ASU D	0.763	62.367	-1.017	218	ASU CD	12.953	30.360	-2.136
218	ASU CG	10.031	30.366	-2.363	218	ASU DD1	14.612	29.700	-3.422
218	ASU DD2	10.660	30.644	-1.103	219	GLV M	0.670	30.934	-3.280
219	GLV CA	0.382	30.130	-2.049	219	GLV C	7.570	37.304	-3.601
219	GLV D	7.073	37.000	-4.076	220	TMR M	6.561	36.638	-3.203
220	TMR CA	3.697	35.936	-4.179	220	TMR C	4.070	37.044	-4.066
220	TMR D	4.617	36.762	-3.050	220	TMR CD	4.825	36.810	-3.926
220	TMR DD1	6.136	38.343	-2.491	220	TMR CD2	3.704	33.094	-2.900
221	SRH M	4.738	30.232	-4.303	221	SRH CA	3.084	39.201	-3.169
221	SRH C	4.760	30.643	-6.303	221	SRH D	4.117	40.208	-7.277
221	SRH CD	3.323	40.383	-4.366	221	SRH DD	3.433	40.282	-3.140
222	RET M	0.065	39.389	-8.485	222	RET CE	6.671	42.771	-3.173
222	RET SD	7.768	41.933	-4.993	222	RET CG	0.804	41.390	-6.602
222	RET CO	0.351	40.018	-7.210	222	RET CA	6.916	39.670	-7.638
222	RET C	6.077	38.435	-8.567	222	RET D	7.004	38.567	-9.775
223	ALA M	6.554	37.246	-8.041	223	ALA CA	6.469	36.020	-8.815
223	ALA C	3.200	36.068	-9.707	223	ALA D	0.133	35.948	-10.929
223	ALA CD	6.909	34.807	-7.923	224	SRH M	4.076	36.360	-9.038
224	SRH CA	3.758	36.489	-9.700	224	SRH C	2.661	37.161	-11.039
224	SRH D	2.163	36.393	-12.057	224	SRH CD	1.002	36.995	-8.603
224	SRH DD	0.492	36.099	-9.137	225	PRD M	3.156	38.411	-11.159
225	PRD CA	3.095	39.130	-12.439	225	PRD C	3.766	38.469	-13.674
225	PRD D	3.406	38.650	-14.804	225	PRD CD	0.653	40.911	-12.054
225	PRD CG	4.611	40.402	-10.764	225	PRD CD	3.735	39.224	-10.094
226	MIS M	4.769	37.626	-13.299	226	MIS CA	3.446	36.070	-14.362
226	MIS C	4.418	35.947	-15.061	226	MIS D	4.425	35.809	-16.293
226	MIS CD	0.608	36.046	-13.765	226	MIS CG	7.814	36.859	-13.338
226	MIS DD1	8.048	37.488	-12.170	226	MIS CD2	0.883	37.110	-14.167
226	MIS CD1	9.270	38.052	-12.236	226	MIS ME2	0.771	37.066	-13.463
227	VAL M	3.593	35.366	-14.199	227	VAL CA	2.983	36.388	-14.727
227	VAL C	3.479	35.197	-15.421	227	VAL D	1.018	36.773	-16.490
227	VAL CD	2.103	33.444	-13.619	227	VAL CD1	1.076	32.476	-14.266
227	VAL CD2	3.204	32.665	-12.891	228	ALA M	1.003	36.242	-14.814
228	ALA CA	0.011	37.109	-15.517	228	ALA C	0.943	37.938	-16.968
228	ALA D	-0.253	37.433	-17.028	228	ALA CD	-0.307	38.353	-16.668
229	GLV M	1.791	38.028	-16.941	229	GLV CA	2.352	38.408	-18.239
229	GLV C	2.420	37.197	-19.187	229	GLV D	2.109	37.375	-20.304
230	ALA M	2.711	38.988	-18.646	230	ALA CA	2.794	34.001	-19.546
230	ALA C	1.424	36.800	-20.133	230	ALA D	1.380	34.203	-21.343
230	ALA CD	3.298	33.624	-18.709	231	ALA M	0.385	34.623	-19.328
231	ALA CA	-1.010	34.416	-19.744	231	ALA C	-1.256	33.423	-20.064
231	ALA D	-1.909	39.056	-21.952	231	ALA CD	-1.032	34.664	-18.549
232	ALA M	-0.778	36.657	-21.721	232	ALA CA	-1.013	37.663	-21.792
232	ALA C	-0.281	37.204	-23.078	232	ALA D	-0.061	37.901	-24.187
232	ALA CD	-0.742	39.121	-21.377	233	LEU M	0.935	36.724	-22.967
233	LEU CA	1.617	36.293	-24.209	233	LEU C	0.821	35.169	-24.880
233	LEU D	0.696	33.231	-20.111	233	LEU CD	3.643	35.877	-23.907
233	LEU CG	3.996	36.994	-23.653	233	LEU CD1	5.239	36.362	-22.921
233	LEU CD2	4.241	37.853	-24.680	234	ILE M	0.357	34.199	-24.047
234	ILE CD1	0.306	30.664	-21.657	234	ILE CG1	0.484	31.223	-23.103
234	ILE CD	-0.011	32.014	-23.570	234	ILE CG2	-1.803	30.900	-24.091
234	ILE CA	-0.406	33.076	-24.644	234	ILE C	-1.621	33.997	-29.434
234	ILE D	-1.083	33.144	-24.544	235	LEU M	-2.390	34.465	-24.779
235	LEU CA	-3.396	35.028	-25.423	235	LEU C	-3.258	39.843	-26.072
235	LEU D	-4.109	35.914	-27.989	235	LEU CD	-4.432	35.765	-24.378
235	LEU CG	-3.140	34.899	-23.340	235	LEU CD1	-3.652	35.603	-22.145
235	LEU CD2	-6.252	34.338	-24.120	236	SRH M	-2.094	34.438	-26.798
236	SRH CA	-1.764	37.237	-27.986	236	SRH C	-1.491	36.292	-29.144
236	SRH D	-1.746	36.634	-30.290	236	SRH CD	-0.633	38.234	-27.733
236	SRH DD	0.599	37.571	-27.582	237	LVS M	-1.044	35.067	-28.002
237	LVS CA	-0.046	34.085	-29.952	237	LVS C	-2.113	33.277	-30.268
237	LVS D	-1.378	32.931	-31.644	237	LVS CD	0.272	33.112	-29.391
237	LVS CG	0.677	32.240	-30.716	237	LVS CD	3.020	31.035	-30.642

237	LVS CE	2.343	30.762	-21.724	237	LVS M2	3.823	20.848	-21.966
238	MIS M	-2.953	31.989	-29.315	238	MIS CA	-4.168	31.163	-29.379
239	MIS C	-5.334	31.899	-29.697	239	MIS D	-5.753	32.984	-27.962
240	MIS CD	-3.968	30.862	-28.811	240	MIS CC	-3.888	29.921	-29.237
241	MIS CD1	-1.707	29.679	-28.833	241	MIS CD2	-3.137	29.238	-29.396
242	MIS CD2	-1.886	29.931	-29.642	242	MIS MF2	-1.949	28.680	-29.399
243	PBD M	-5.848	33.917	-29.363	243	PBD CA	-6.988	34.779	-28.773
244	PBD C	-8.204	34.252	-28.932	244	PBD D	-8.949	34.919	-27.462
245	PBD CB	-7.818	35.977	-29.713	245	PBD CC	-6.666	35.294	-28.827
246	PBD CD	-5.436	34.639	-29.668	246	ASN M	-3.386	32.969	-29.227
247	ASN CA	-9.829	37.043	-29.216	247	ASN C	-9.593	31.180	-27.988
248	ASN D	-10.340	38.610	-27.576	248	ASN CB	-9.403	31.249	-29.535
249	ASN CC	-7.971	38.827	-29.889	249	ASN CD1	-7.898	31.590	-29.147
250	ASN CD2	-7.675	29.909	-28.976	250	TRP M	-8.354	31.806	-27.384
251	TRP CA	-8.304	30.124	-26.126	251	TRP C	-9.106	30.638	-26.936
252	TRP D	-9.843	31.833	-24.686	252	TRP CB	-6.879	29.835	-25.679
253	TRP CC	-6.094	28.903	-26.937	253	TRP CD1	-6.338	28.433	-27.818
254	TRP CD2	-6.839	28.374	-26.193	254	TRP ME1	-3.362	27.547	-28.211
255	TRP CE2	-6.414	27.674	-27.216	255	TRP CE3	-4.097	28.496	-24.981
256	TRP CE2	-3.193	26.786	-27.174	256	TRP CE3	-2.912	27.667	-24.943
257	TRP CM2	-2.470	26.873	-26.093	257	TMR M	-9.727	29.781	-24.162
258	TMR CA	-10.498	38.119	-22.911	258	TMR C	-9.469	30.176	-21.767
259	TMR D	-9.335	29.674	-21.937	259	TMR CB	-11.579	29.832	-22.675
260	TMR DG1	-10.837	27.786	-22.476	260	TMR CC2	-12.494	28.907	-23.898
261	ASN M	-9.946	30.639	-20.611	261	ASN MD2	-11.787	28.484	-23.767
262	ASN CD1	-11.463	31.918	-18.758	262	ASN CC	-11.893	31.131	-17.985
263	ASN CB	-9.701	31.930	-18.332	263	ASN CA	-9.853	30.731	-19.446
264	ASN C	-8.637	29.383	-19.810	264	ASN D	-7.593	29.136	-18.460
265	TMR M	-9.364	28.362	-19.233	265	TMR CA	-9.381	26.934	-19.859
266	TMR C	-8.133	26.393	-19.882	266	TMR D	-7.324	25.757	-19.111
267	TMR CB	-10.463	24.888	-19.494	267	TMR DG1	-11.735	24.675	-18.684
268	TMR CC2	-10.553	24.593	-19.158	268	GLN M	-8.882	24.716	-21.073
269	GLN CA	-6.964	26.362	-21.962	269	GLN C	-8.647	27.020	-21.820
270	GLN D	-6.573	26.393	-21.447	270	GLN CB	-7.330	24.999	-23.397
271	GLN CC	-8.265	25.526	-23.989	271	GLN CD	-8.493	29.873	-23.428
272	GLN CD1	-9.356	26.769	-23.727	272	GLN ME2	-7.745	25.312	-23.370
273	VAL M	-5.697	28.304	-21.218	273	VAL CA	-4.477	29.040	-20.778
274	VAL C	-3.936	26.462	-19.467	274	VAL D	-2.785	28.227	-19.361
275	VAL CB	-4.779	30.553	-20.623	275	VAL CC1	-3.564	31.272	-20.827
276	VAL CC2	-5.169	31.138	-21.959	276	ARC M	-4.767	28.240	-18.462
277	ARC CA	-6.388	27.714	-17.168	277	ARC C	-3.770	26.292	-17.360
278	ARC D	-2.701	25.985	-18.764	278	ARC CB	-5.533	27.667	-16.149
279	ARC CC	-4.987	27.093	-14.892	279	ARC CD	-6.896	27.179	-13.793
280	ARC ME	-5.440	26.757	-12.846	280	ARC CI	-3.893	26.866	-11.313
281	ARC MD1	-7.064	27.484	-11.210	281	ARC MD2	-5.177	24.428	-10.270
282	SER M	-4.480	25.809	-18.131	282	SER CA	-6.839	24.131	-18.426
283	SER C	-2.637	24.886	-19.073	283	SER D	-1.848	23.253	-18.583
284	SER CB	-5.034	23.409	-19.372	284	SER DG	-6.146	23.896	-18.832
285	SER M	-2.500	24.883	-20.136	285	SER CA	-1.223	24.874	-20.851
286	SER C	-0.071	25.302	-19.940	286	SER D	1.826	24.788	-20.049
287	SER CB	-1.369	25.798	-22.068	287	SER DG	-3.300	25.619	-22.956
288	LEU M	-0.289	26.333	-19.160	288	LEU CD2	1.824	29.814	-18.222
289	LEU CD1	-0.373	28.433	-17.268	289	LEU CC	0.352	29.438	-18.151
290	LEU CB	0.178	28.083	-17.563	290	LEU CA	0.718	24.837	-18.216
291	LEU C	1.092	25.694	-17.263	291	LEU C	2.293	25.421	-17.032
292	GLN M	0.868	25.827	-16.714	292	GLN ME2	-2.780	25.912	-12.237
293	GLN DG1	-2.819	23.624	-12.933	293	GLN CD	-2.943	24.910	-13.834
294	GLN CC	-1.210	24.814	-13.994	294	GLN CB	-8.897	25.621	-14.977
295	GLN CA	0.381	22.941	-13.745	295	GLN C	0.919	22.664	-16.361
296	GLN D	1.743	22.014	-15.616	296	ASN M	0.633	22.394	-17.998
297	ASN CA	1.882	21.201	-18.282	297	ASN C	2.394	21.359	-18.991
298	ASN D	2.899	20.442	-19.760	298	ASN CB	0.894	20.780	-19.292
299	ASN CC	-1.036	19.826	-18.573	299	ASN CD1	-8.836	19.385	-17.582

252	AS4 MD2	-2.234	29.874	-19.541	253	TM2 M	9.818	22.585	-18.923
253	TM2 CA	4.256	22.717	-19.713	253	TM2 E	9.361	23.247	-18.818
253	TM2 D	0.368	23.793	-19.427	253	TM2 CB	4.086	23.672	-18.952
253	TM2 DC1	3.993	24.037	-20.428	253	TM2 CC2	3.147	23.139	-22.031
254	TM2 M	8.218	23.177	-17.851	254	TM2 CA	6.214	23.612	-18.588
254	TM2 C	7.466	22.750	-18.612	254	TM2 D	7.402	21.950	-17.893
254	TM2 CB	9.664	23.938	-19.132	254	TM2 DC1	9.129	22.178	-18.040
254	TM2 CC2	4.535	24.549	-14.802	255	TM2 M	8.499	23.296	-14.876
255	TM2 CA	9.771	22.544	-15.817	255	TM2 E	9.671	22.031	-14.414
255	TM2 D	9.439	22.786	-13.474	255	TM2 CB	11.080	23.453	-15.897
255	TM2 DC1	11.892	23.789	-17.321	255	TM2 CC2	12.284	22.628	-19.404
256	LVS M	9.896	20.782	-14.314	256	LVS CA	9.344	20.063	-13.810
256	LVS C	10.322	20.323	-12.063	256	LVS D	11.662	20.274	-12.992
256	LVS CB	9.024	18.990	-13.249	256	LVS CC	9.818	17.805	-11.921
256	LVS CD	10.284	16.948	-11.777	256	LVS CE	10.212	19.940	-18.623
256	LVS CZ	9.243	14.969	-11.054	257	LEU M	10.212	20.474	-18.874
257	LEU CA	11.272	21.035	-9.893	257	LEU C	11.250	20.232	-8.614
257	LEU D	12.094	20.845	-7.732	257	LEU CB	11.187	22.947	-9.522
257	LEU CC	11.397	23.420	-10.848	257	LEU CD1	11.245	21.803	-9.921
257	LEU CC2	12.678	23.468	-11.325	258	GLY M	10.431	19.282	-8.298
258	GLY CA	10.602	18.783	-6.879	258	GLY C	9.168	18.703	-4.373
258	GLY D	8.283	18.936	-7.202	259	ASP M	9.824	18.232	-5.190
259	ASP CA	7.757	17.896	-4.516	259	ASP C	6.639	18.941	-4.789
259	ASP D	6.859	20.039	-4.214	259	ASP CB	7.994	17.540	-3.833
259	ASP CC	4.781	17.128	-2.241	259	ASP CD1	5.611	17.527	-2.354
259	ASP CD2	7.998	16.299	-1.321	260	SER M	5.960	18.610	-5.312
260	SER CA	4.481	19.587	-5.529	260	SER C	4.046	20.362	-6.289
260	SER D	3.500	21.553	-4.446	260	SER CB	3.945	18.919	-4.289
260	SER CC	2.745	17.937	-3.440	261	PHE M	4.241	19.778	-3.112
261	PHE CA	3.831	28.468	-1.845	261	PHE C	4.544	21.846	-1.863
261	PHE D	3.944	22.848	-1.432	261	PHE CB	4.053	19.749	-0.563
261	PHE CC	3.949	20.337	0.719	261	PHE CD1	2.204	20.163	1.123
261	PHE CD2	4.401	21.060	1.538	261	PHE CE1	1.737	20.717	2.315
261	PHE CE2	3.945	21.802	2.748	261	PHE C2	2.605	21.465	3.114
262	TYR M	5.778	21.788	-2.305	262	TYR CA	4.688	22.914	-2.251
262	TYR C	6.820	23.689	-3.545	262	TYR D	7.201	24.853	-3.393
262	TYR CB	8.122	22.455	-1.831	262	TYR CC	8.146	21.892	-0.454
262	TYR CD1	9.084	20.484	-0.344	262	TYR CD2	8.149	22.669	0.498
262	TYR CE1	8.062	19.873	0.882	262	TYR CE2	8.314	22.669	1.962
262	TYR C2	8.069	20.672	2.010	262	TYR DM	7.945	20.829	3.205
263	TYR M	6.616	23.104	-4.693	263	TYR CA	6.812	23.659	-6.022
263	TYR C	5.624	23.680	-6.954	263	TYR D	5.781	24.117	-5.111
263	TYR CB	7.928	22.768	-6.681	263	TYR CC	9.279	23.035	-4.048
263	TYR CD1	10.044	24.046	-6.637	263	TYR CD2	9.800	22.542	-4.995
263	TYR CE1	11.395	24.328	-6.168	263	TYR CE2	11.062	22.640	-6.491
263	TYR C2	11.838	23.618	-5.106	263	TYR DM	17.065	23.949	-4.597
264	GLY M	4.473	23.161	-6.516	264	GLY CA	3.301	23.064	-7.412
264	GLY C	3.847	22.196	-8.554	264	GLY D	4.447	21.274	-8.365
265	LVS M	3.436	22.477	-9.754	265	LVS CA	3.834	21.788	-10.971
265	LVS C	9.188	22.232	-11.464	265	LVS D	5.684	21.863	-12.384
265	LVS CB	2.755	22.071	-12.044	265	LVS CC	1.490	21.543	-11.308
265	LVS CD	9.710	20.848	-12.079	265	LVS CE	-8.692	20.496	-11.991
265	LVS CZ	-1.678	23.757	-12.489	266	GLY M	5.787	23.226	-10.817
266	GLY CA	7.120	23.612	-11.323	266	GLY C	7.193	23.052	-11.818
266	GLY D	6.177	23.793	-11.648	267	LEU M	5.262	25.354	-12.480
267	LEU CA	8.490	26.460	-13.097	267	LEU C	7.804	26.771	-14.437
267	LEU D	7.953	25.909	-15.298	267	LEU CB	10.010	26.955	-13.216
267	LEU CC	10.432	28.050	-14.058	267	LEU CD1	10.994	29.331	-13.250
267	LEU CD2	11.924	27.921	-14.927	268	ILE M	7.064	27.063	-14.632
268	ILE CA	4.406	28.035	-15.944	268	ILE C	7.426	28.246	-17.065
268	ILE D	8.939	28.793	-16.912	268	ILE CB	8.369	29.210	-15.899
268	ILE CC1	6.099	30.541	-15.552	268	ILE CC2	4.243	28.923	-14.861
268	ILE CD1	5.389	31.785	-16.242	269	ASN M	7.987	27.843	-18.237

269	ALA CA	1.362	27.973	-20.497	269	ALA C	0.939	28.934	-20.498
269	ALA D	5.963	27.762	-20.602	269	ALA CB	0.457	28.613	-19.895
269	ALA CL	9.163	26.856	-21.210	269	ALA CD1	0.993	27.424	-21.122
269	ALA MD2	12.033	25.796	-21.672	270	VAL H	0.908	29.868	-20.720
270	VAL CA	2.063	26.418	-21.014	270	VAL C	0.899	29.907	-21.684
270	VAL D	2.067	25.969	-21.872	270	VAL CB	0.496	28.950	-21.422
270	VAL CL1	0.969	22.797	-21.979	270	VAL CD2	0.429	28.362	-22.292
271	GLN H	5.323	29.703	-21.792	271	GLN CA	7.603	27.370	-26.764
271	GLN C	6.869	27.934	-21.031	271	GLN D	0.213	27.204	-26.091
271	GLN CB	0.104	25.220	-21.964	271	GLN CC	0.486	28.610	-26.336
271	GLN CD	10.901	28.513	-21.892	271	GLN D13	11.269	28.979	-27.713
271	GLN MD2	11.702	28.333	-21.910	272	ALA H	0.977	28.909	-24.082
272	ALA CA	0.224	25.712	-24.140	272	ALA C	0.701	28.958	-24.266
272	ALA D	3.973	23.303	-23.002	272	ALA CB	0.763	28.742	-22.172
272	ALA H	0.267	26.053	-23.239	272	ALA CA	2.640	28.723	-22.694
272	ALA C	2.081	27.329	-24.020	272	ALA D	0.999	27.219	-24.285
272	ALA CB	2.736	27.773	-21.995	274	ALA H	1.789	28.464	-24.742
274	ALA CB	2.932	30.391	-20.210	274	ALA CA	2.189	29.164	-25.067
274	ALA C	1.730	27.367	-27.090	274	ALA D	0.980	28.949	-27.021
275	GLN H	2.820	27.144	-27.714	275	GLN CA	2.848	28.309	-28.827
275	GLN C	1.147	27.261	-27.777	275	GLN D	2.260	27.007	-27.916
275	GLN D1	3.193	27.363	-20.590	275	GLN CB	0.634	28.724	-28.820
275	GLN CC	0.571	26.604	-27.447	275	GLN CD	-0.023	28.986	-27.631
275	GLN D11	-1.376	23.895	-20.729	275	GLN MD2	-1.373	28.411	-26.538

The above structural studies together with the kinetic data presented herein and elsewhere (Philipp, M., et al. (1983) Mol. Cell. Biochem. 51, 5-32; Svendsen, I.B. (1976) Carlsberg Res. Comm. 41, 237-291; Markland, S.F. Id; Stauffe, D.C., et al. (1965) J. Biol. Chem. 244, 5333-5338) indicate that the subsites in the binding cleft of subtilisin are capable of interacting with substrate amino acid residues from P-4 to P-2'.

10 The most extensively studied of the above residues are Gly166, Gly169 and Ala152. These amino acids were identified as residues within the S-1 subsite. As seen in Fig. 3, which is a stereoview of the S-1 subsite, Gly166 and Gly169 occupy positions at the bottom of the S-1 subsite, whereas Ala152 occupies a position near the top of S-1, close to the catalytic Ser221.

20 All 19 amino acid substitutions of Gly166 and Gly169 have been made. As will be indicated in the examples which follow, the preferred replacement amino acids for Gly166 and/or Gly169 will depend on the specific amino acid occupying the P-1 position of a given substrate.

25 The only substitutions of Ala152 presently made and analyzed comprise the replacement of Ala152 with Gly and Ser. The results of these substitutions on P-1 specificity will be presented in the examples.

30 In addition to those residues specifically associated with specificity for the P-1 substrate amino acid, Tyr104 has been identified as being involved with P-4 specificity. Substitutions at Phe189 and Tyr217,

35

however, are expected to respectively effect P-2' and P-1' specificity.

5 The catalytic activity of subtilisin has also been modified by single amino acid substitutions at Asn155. The catalytic triad of subtilisin is shown in Fig. 4. As can be seen, Ser221, His64 and Asp32 are positioned to facilitate nucleophilic attack by the serine hydroxylate on the carbonyl of the scissile peptide bond. Crystallographic studies of subtilisin  
10 (Robertus, *et al.* (1972) Biochem. 11, 4293-4303; Matthews, *et al.* (1975) J. Biol. Chem. 250, 7120-7126; Poulos, *et al.* (1976) J. Biol. Chem. 250, 1097-1103) show that two hydrogen bonds are formed with the oxyanion of the substrate transition state. One  
15 hydrogen bond donor is from the catalytic serine-221 main-chain amide while the other is from one of the NE2 protons of the asparagine-155 side chain. See Fig. 4.

20 Asn155 was substituted with Ala, Asp, His, Glu and Thr. These substitutions were made to investigate the the stabilization of the charged tetrahedral intermediate of the transition state complex by the potential hydrogen bond between the side chain of  
25 Asn155 and the oxyanion of the intermediate. These particular substitutions caused large decreases in substrate turnover,  $k_{cat}$  (200 to 4,000 fold), marginal decreases in substrate binding  $K_m$  (up to 7 fold), and  
30 a loss in transition state stabilization energy of 2.2 to 4.7 kcal/mol. The retention of  $K_m$  and the drop in  $k_{cat}$  will make these mutant enzymes useful as binding proteins for specific peptide sequences, the nature of which will be determined by the specificity of the  
35 precursor protease.



Various other amino acid residues have been identified which affect alkaline stability. In some cases, mutants having altered alkaline stability also have altered thermal stability.

- 5 In *B. amyloliquefaciens* subtilisin residues Asp36, Ile107, Lys170, Ser204 and Lys213 have been identified as residues which upon substitution with a different amino acid alter the alkaline stability of the mutated enzyme as compared to the precursor enzyme. The  
10 substitution of Asp36 with Ala and the substitution of Lys170 with Glu each resulted in a mutant enzyme having a lower alkaline stability as compared to the wild type subtilisin. When Ile107 was substituted with Val, Ser204 substituted with Cys, Arg or Leu or  
15 Lys213 substituted with Arg, the mutant subtilisin had a greater alkaline stability as compared to the wild type subtilisin. However, the mutant Ser204P demonstrated a decrease in alkaline stability.
- 20 In addition, other residues, identified as being associated with the modification of other properties of subtilisin, also affect alkaline stability. These residues include Ser24, Met50, Glu156, Gly166, Gly169 and Tyr217. Specifically the following particular  
25 substitutions result in an increased alkaline stability: Ser24C, Met50F, Gly156Q or S, Gly166A, H, K, N or Q, Gly169S or A, and Tyr217F, K, R or L. The mutant Met50V, on the other hand, results in a decrease in the alkaline stability of the mutant  
30 subtilisin as compared to wild type subtilisin.

Other residues involved in alkaline stability based on the alkaline stability screen include Asp197 and Met222. Particular mutants include Asp197(R or A) and  
35 Met 222 (all other amino acids).

Various other residues have been identified as being involved in thermal stability as determined by the thermal stability screen herein. These residues include the above identified residues which effect alkaline stability and Met199 and Tyr21. These latter two residues are also believed to be important for alkaline stability. Mutants at these residues include I199 and F21.

The amino acid sequence of B. amyloliquefaciens subtilisin has also been modified by substituting two or more amino acids of the wild-type sequence. Six categories of multiply substituted mutant subtilisin have been identified. The first two categories comprise thermally and oxidatively stable mutants. The next three other categories comprise mutants which combine the useful properties of any of several single mutations of B. amyloliquefaciens subtilisin. The last category comprises mutants which have modified alkaline and/or thermal stability.

The first category comprises double mutants in which two cysteine residues have been substituted at various amino acid residue positions within the subtilisin molecule. Formation of disulfide bridges between the two substituted cysteine residues results in mutant subtilisins with altered thermal stability and catalytic activity. These mutants include A21/C22/C87 and C24/C87 which will be described in more detail in Example 11.

The second category of multiple subtilisin mutants comprises mutants which are stable in the presence of various oxidizing agents such as hydrogen peroxide or peracids. Examples 1 and 2 describe these mutants

-50-

which include F50/I124/Q222, F50/I124, F50/Q222, F50/L124/Q222, I124/Q222 and L124/Q222.

5 The third category of multiple subtilisin mutants comprises mutants with substitutions at position 222 combined with various substitutions at positions 166 or 169. These mutants, for example, combine the property of oxidative stability of the A222 mutation with the altered substrate specificity of the various 10 166 or 169 substitutions. Such multiple mutants include A166/A222, A166/C222, F166/C222, K166/A222, K166/C222, V166/A222 and V166/C222. The K166/A222 mutant subtilisin, for example, has a kcat/Km ratio which is approximately two times greater than that of 15 the single A222 mutant subtilisin when compared using a substrate with phenylalanine as the P-1 amino acid. This category of multiple mutant is described in more detail in Example 12.

20 The fourth category of multiple mutants combines substitutions at position 156 (Glu to Q or S) with the substitution of Lys at position 166. Either of these single mutations improve enzyme performance upon substrates with glutamate as the P-1 amino acid. When 25 these single mutations are combined, the resulting multiple enzyme mutants perform better than either precursor. See Example 9.

30 The fifth category of multiple mutants contain the substitution of up to four amino acids of the B. amyloliquefaciens subtilisin sequence. These mutants have specific properties which are virtually identical to the properties of the subtilisin from B. licheniformis. The subtilisin from B. licheniformis differs from B. amyloliquefaciens subtilisin at 87 out 35 of 275 amino acids. The multiple mutant

F50/S156/A169/L217 was found to have similar substrate specificity and kinetics to the licheniformis enzyme. (See Example 13.) However, this is probably due to only three of the mutations (S156, A169 and L217) which are present in the substrate binding region of the enzyme. It is quite surprising that, by making only three changes out of the 87 different amino acids between the sequence of the two enzymes, the B. amyloliquifaciens enzyme was converted into an enzyme with properties similar to B. licheniformis enzyme. Other enzymes in this series include F50/Q156/N166/L217 and F50/S156/L217.

The sixth category of multiple mutants includes the combination of substitutions at position 107 (Ile to V) with the substitution of Lys at position 213 with Arg, and the combination of substitutions of position 204 (preferably Ser to C or L but also to all other amino acids) with the substitution of Lys at position 213 with R. Other multiple mutants which have altered alkaline stability include Q156/K166, Q156/N166, S156/K166, S156/N166 (previously identified as having altered substrate specificity), and F50/S156/A169/L217 (previously identified as a mutant of B. amyloliquifaciens subtilisin having properties similar to subtilisin from B. licheniformis). The mutant F50/V107/R213 was constructed based on the observed increase in alkaline stability for the single mutants F50, V107 and R213. It was determined that the V107/R213 mutant had an increased alkaline stability as compared to the wild type subtilisin. In this particular mutant, the increased alkaline stability was the result of the cumulative stability of each of the individual mutations. Similarly, the mutant F50/V107/R213 had an even greater alkaline stability as compared to the V107/R213 mutant indicating that

the increase in the alkaline stability due to the F50 mutation was also cumulative.

Table IV summarizes the multiple mutants which have been made including those not mentioned above.

5

In addition, based in part on the above results, substitution at the following residues in subtilisin is expected to produce a multiple mutant having increased thermal and alkaline stability: Ser24,  
10 Met50, Ile107, Glu156, Gly166, Gly169, Ser204, Lys213, Gly215, and Tyr217.

15

20

25

30

35

TABLE IV

	<u>Double Mutants</u>	<u>Triple, Quadruple or Other Multiple</u>
	C22/C87	F50/I124/Q222
	C24/C87	F50/L124/Q222
5	V45/V48	F50/L124/A222
	C49/C94	A21/C22/C87
	C49/C95	F50/S156/N166/L217
	C50/C95	F50/Q156/N166/L217
	C50/C110	F50/S156/A169/L217
10	F50/I124	F50/S156/L217
	F50/Q222	F50/Q156/K166/L217
	I124/Q222	F50/S156/K166/L217
	Q156/D166	F50/Q156/K166/K217
	Q156/K166	F50/S156/K166/K217
15	Q156/N166	F50/V107/R213
	S156/D166	[S153/S156/A158/G159/S160/A161- 164/I165/S166/A169/R170]
	S156/K166	L204/R213
	S156/N166	L204/R213
	S156/A169	R213/204A, E, Q, D, N, G, K, V, R, T, P, I, M, F, Y, W or H
20	A166/A222	
	A166/C222	
	F166/A222	V107/R213
	F166/C222	
	K166/A222	
25	K166/C222	
	V166/A222	
	V166/C222	
	A169/A222	
	A169/A222	
30	A169/C222	
	A21/C22	

In addition to the above identified amino acid  
 35 residues, other amino acid residues of subtilisin are

also considered to be important with regard to substrate specificity. Mutation of each of these residues is expected to produce changes in the substrate specificity of subtilisin. Moreover, multiple mutations among these residues and among the previously identified residues are also expected to produce subtilisin mutants having novel substrate specificity.

Particularly important residues are His67, Ile107, Leu126 and Leu135. Mutation of His67 should alter the S-1' subsite, thereby altering the specificity of the mutant for the P-1' substrate residue. Changes at this position could also affect the pH activity profile of the mutant. This residue was identified based on the inventor's substrate modeling from product inhibitor complexes.

Ile107 is involved in P-4 binding. Mutation at this position thus should alter specificity for the P-4 substrate residue in addition to the observed effect on alkaline stability. Ile107 was also identified by molecular modeling from product inhibitor complexes.

The S-2 binding site includes the Leu126 residue. Modification at this position should therefore affect P-2 specificity. Moreover, this residue is believed to be important to convert subtilisin to an amino peptidase. The pH activity profile should also be modified by appropriate substitution. These residues were identified from inspection of the refined model, the three dimensional structure from modeling studies. A longer side chain is expected to preclude binding of any side chain at the S-2 subsite. Therefore, binding would be restricted to subsites S-1, S-1', S-2', S-3'

and cleavage would be forced to occur after the amino terminal peptide.

Leu135 is in the S-4 subsite and if mutated should alter substrate specificity for P-4 if mutated. This residue was identified by inspection of the three-dimensional structure and modeling based on the product inhibitor complex of F222.

In addition to these sites, specific amino acid residues within the segments 97-103, 126-129 and 213-215 are also believed to be important to substrate binding.

Segments 97-103 and 126-129 form an antiparallel beta sheet with the main chain of substrate residues P-4 through P-2. Mutating residues in those regions should affect the substrate orientation through main chain (enzyme) - main chain (substrate) interactions, since the main chain of these substrate residues do not interact with these particular residues within the S-4 through S-2 subsites.

Within the segment 97-103, Gly97 and Asp99 may be mutated to alter the position of residues 101-103 within the segment. Changes at these sites must be compatible, however. In B. amyloliquefaciens subtilisin Asp99 stabilizes a turn in the main chain tertiary folding that affects the direction of residues 101-103. B. licheniformis subtilisin Asp97, functions in an analogous manner.

In addition to Gly97 and Asp99, Ser101 interacts with Asp99 in B. amyloliquefaciens subtilisin to stabilize the same main chain turn. Alterations at this residue should alter the 101-103 main chain direction.



Mutations at Glu103 are also expected to affect the 101-103 main chain direction.

The side chain of Gly102 interacts with the substrate P-3 amino acid. Side chains of substituted amino acids thus are expected to significantly affect specificity for the P-3 substrate amino acids.

All the amino acids within the 127-129 segment are considered important to substrate specificity. Gly 127 is positioned such that its side chain interacts with the S-1 and S-3 subsites. Altering this residue thus should alter the specificity for P-1 and P-3 residues of the substrate.

The side chain of Gly128 comprises a part of both the S-2 and S-4 subsites. Altered specificity for P-2 and P-4 therefore would be expected upon mutation. Moreover, such mutation may convert subtilisin into an amino peptidase for the same reasons substitutions of Leu126 would be expected to produce that result.

The Pro129 residue is likely to restrict the conformational freedom of the sequence 126-133, residues which may play a major role in determining P-1 specificity. Replacing Pro may introduce more flexibility thereby broadening the range of binding capabilities of such mutants.

The side chain of Lys213 is located within the S-3 subsite. All of the amino acids within the 213-215 segment are also considered to be important to substrate specificity. Accordingly, altered P-3 substrate specificity is expected upon mutation of this residue.

The Tyr214 residue does not interact with substrate but is positioned such that it could affect the conformation of the hair pin loop 204-217.

5 Finally, mutation of the Gly215 residue should affect the S-3' subsite, and thereby alter P-3' specificity.

10 In addition to the above substitutions of amino acids, the insertion or deletion of one or more amino acids within the external loop comprising residues 152-172 may also affect specificity. This is because these residues may play a role in the "secondary contact region" described in the model of streptomyces subtilisin inhibitor complexed with subtilisin. Hirono, et al. (1984) J. Mol. Biol. 178, 389-413.

15 Thermitase K has a deletion in this region, which eliminates several of these "secondary contact" residues. In particular, deletion of residues 161 through 164 is expected to produce a mutant subtilisin having modified substrate specificity. In addition, a

20 rearrangement in this area induced by the deletion should alter the position of many residues involved in substrate binding, predominantly at P-1. This, in turn, should affect overall activity against proteinaceous substrates.

25

The effect of deletion of residues 161 through 164 has been shown by comparing the activity of the wild type (WT) enzyme with a mutant enzyme containing this deletion as well as multiple substitutions (i.e.,

30 S153/S156/A158/G159/S160/ $\Delta$ 161-164/I165/S166/A169/R170). This produced the following results:

35

TABLE V

	<u>kcat</u>	<u>Km</u>	<u>kcat/Km</u>
WT	50	$1.4 \times 10^{-4}$	$3.6 \times 10^5$
Deletion mutant	8	$5.0 \times 10^{-6}$	$1.6 \times 10^6$

5

The WT has a kcat 6 times greater than the deletion mutant but substrate binding is 28 fold tighter by the deletion mutant. The overall efficiency of the deletion mutant is thus 4.4 times higher than the WT enzyme.

All of these above identified residues which have yet to be substituted, deleted or inserted into are presented in Table VI.

15

TABLE VI

Substitution/Insertion/Deletion

20

Residues

25

30

35

His67	Ala152
Leu126	Ala153
Leu135	Gly154
Gly97	Asn155
Asp99	Gly156
Ser101	Gly157
Gly102	Gly160
Glu103	Thr158
Leu126	Ser159
Gly127	Ser161
Gly128	Ser162
Pro129	Ser163
Tyr214	Thr164
Gly215	Val165
Gly166	Gly169
Tyr167	Lys170
Pro168	Tyr171
	Pro172

The following disclosure is intended to serve as a representation of embodiments herein, and should not be construed as limiting the scope of this application. These specific examples disclose the construction of certain of the above identified mutants. The construction of the other mutants, however, is apparent from the disclosure herein and that presented in EPO Publication No. 0130756.

All literature citations are expressly incorporated by reference.

EXAMPLE 1

Identification of Peracid Oxidizable Residues of Subtilisin Q222 and L222

As shown in Figures 6A and 6B, organic peracid oxidants inactivate the mutant subtilisins Met222L and Met222Q (L222 and Q222). This example describes the identification of peracid oxidizable sites in these mutant subtilisins.

First, the type of amino acid involved in peracid oxidation was determined. Except under drastic conditions (Means, G.E., et al. (1971) Chemical Modifications of Proteins, Holden-Day, S.F., CA, pp. 160-162), organic peracids modify only methionine and tryptophan in subtilisin. Difference spectra of the enzyme over the 250nm to 350nm range were determined during an inactivation titration employing the reagent, diperdodecanoic acid (DPDA) as oxidant. Despite quantitative inactivation of the enzyme, no change in absorbance over this wavelength range was noted as shown in Figures 7A and 7B indicating that tryptophan was not oxidized. Fontana, A., et al. (1980) Methods in Peptide and Protein Sequence

Analysis (C. Birr ed.) Elsevier, New York, p. 309. The absence of tryptophan modification implied oxidation of one or more of the remaining methionines of B. amyloliquefaciens subtilisin. See Figure 1.

5 To confirm this result the recombinant subtilisin Met222F was cleaved with cyanogen bromide (CNBr) both before and after oxidation by DPDA. The peptides produced by CNBr cleavage were analyzed on high resolution SDS-pyridine peptide gels (SPG).

10

Subtilisin Met222F (F222) was oxidized in the following manner. Purified F222 was resuspended in 0.1 M sodium borate pH 9.5 at 10 mg/ml and was added to a final concentration of 26 dperdodecanoic acid (DPDA) at 26 mg/ml was added to produce an effective active oxygen concentration of 30 ppm. The sample was incubated for at least 30 minutes at room temperature and then quenched with 0.1 volume of 1 M Tris pH 8.6 buffer to produce a final concentration of 0.1 M Tris pH 8.6). 3mM phenylmethylsulfonyl fluoride (PMSF) was added and 2.5 ml of the sample was applied to a Pharmacia PD10 column equilibrated in 10 mM sodium phosphate pH 6.2, 1 mM PMSF. 3.5 ml of 10 mM sodium phosphate pH6.2, 1mM PMSF was applied and the eluant collected.

25

F222 and DPDA oxidized F222 were precipitated with 9 volumes of acetone at -20°C. The samples were resuspended at 10 mg/ml in 8M urea in 88% formic acid and allowed to sit for 5 minutes. An equal volume of 200 mg/ml CNBr in 88% formic acid was added (5 mg/ml protein) and the samples incubated for 2 hours at room temperature in the dark. Prior to gel electrophoresis, the samples were lyophilized and resuspended at 2-5 mg/ml in sample buffer (1%

30

35

pyridine, 5% NaDodSO<sub>4</sub>, 5% glycerol and bromophenol blue) and disassociated at 95°C for 3 minutes.

5 The samples were electrophoresed on discontinuous polyacrylamide gels (Kyte, J., et al. (1983) Anal. Bioch. 133, 515-522). The gels were stained using the Pharmacia silver staining technique (Sammons, D.R., et al. (1981) Electrophoresis 2 135-141).

10 The results of this experiment are shown in Figure 8. As can be seen, F222 treated with CNBr only gives nine resolved bands on SPG. However, when F222 is also treated with DPDA prior to cleavage, bands X, 7 and 9 disappear whereas bands 5 and 6 are greatly increased  
15 in intensity.

In order to determine which of the methionines were effected, each of the CNBr peptides was isolated by  
20 reversed phase HPLC and further characterized. The buffer system in both Solvent A (aqueous) and Solvent B (organic) for all HPLC separations was 0.05% triethylamine/trifluoroacetic acid (TEA-TFA). In all cases unless noted, solvent A consisted of 0.05%  
25 TEA-TFA in H<sub>2</sub>O, solvent B was 0.05% TEA-TFA in 1-propanol, and the flow rate was 0.5 ml/minute.

For HPLC analysis, two injections of 1 mg enzyme digest were used. Three samples were acetone  
30 precipitated, washed and dried. The dried 1 mg samples were resuspended at 10 mg/ml in 8M urea, 88% formic acid; an equal volume of 200 mg/ml CNBr in 88% formic acid was added (5 mg/ml protein). After incubation for 2 hours in the dark at room  
35 temperature, the samples were desalted on a 0.8 cm X 7

cm column of Tris Acryl GF05 coarse resin (IBF, Paris, France) equilibrated with 40% solvent B, 60% solvent A. 200 ul samples were applied at a flow rate of 1 ml a minute and 1.0-1.2 ml collected by monitoring the absorbance at 280nm. Prior to injection on the HPLC, each desalted sample was diluted with 3 volumes of solvent A. The samples were injected at 1.0 ml/min (2 minutes) and the flow then adjusted to 0.5 ml/min (100% A). After 2 minutes, a linear gradient to 60% B at 1.0% B/min was initiated. From each 1 mg run, the pooled peaks were sampled (50ul) and analyzed by gel electrophoresis as described above.

Each polypeptide isolated by reversed phase HPLC was further analyzed for homogeneity by SPG. The position of each peptide on the known gene sequence (Wells, J.A., et al. (1983) Nucleic Acids Res. 11 7911-7924) was obtained through a combination of amino acid compositional analysis and, where needed, amino terminal sequencing.

Prior to such analysis the following peptides were to rechromatographed.

1. CNBr peptides from F222 not treated with DPDA:

Peptide 5 was subjected to two additional reversed phase separations. The 10 cm C4 column was equilibrated to 80%A/ 20%B and the pooled sample applied and washed for 2 minutes. Next an 0.5% ml B/min gradient was initiated. Fractions from this separation were again rerun, this time on the 25 cm C4 column, and employing 0.05% TEA-TFA in acetonitrile/1-propanol (1:1) for solvent B. The gradient was identical to the one just described.

Peptide "X" was subjected to one additional separation after the initial chromatography. The sample was applied and washed for 2 minutes at 0.5ml/min (100%A), and a 0.5% ml B/min gradient was initiated.

5 Peptides 7 and 9 were rechromatographed in a similar manner to the first rerun of peptide 5.

Peptide 8 was purified to homogeneity after the initial separation.

10 2. CNBr Peptides from DPDA Oxidized F222:

Peptides 5 and 6 from a CNBr digest of the oxidized F222 were purified in the same manner as peptide 5 from the untreated enzyme.

15 Amino acid compositional analysis was obtained as follows. Samples (~1mM each amino acid) were dried, hydrolyzed in vacuo with 100 ul 6N HCl at 106°C for 24 hours and then dried in a Speed Vac. The samples were  
20 analyzed on a Beckmann 6300 AA analyzer employing ninhydrin detection.

Amino terminal sequence data was obtained as previously described (Rodriguez, H., et al. (1984)  
25 Anal. Biochem. 134, 538-547).

The results are shown in Table VII and Figure 9.

30

35



TABLE VII

Amino and COOH termini of CNBr fragments

<u>Terminus and Method</u>			
	<u>Fragment</u>	<u>amino, method</u>	<u>COOH, method</u>
5	X	1, sequence	50, composition
	9	51, sequence	119, composition
	7	125, sequence	199, composition
	8		
		200, sequence	275, composition
10	5ox	1, sequence	119, composition
	6ox	120, composition	199, composition

15 Peptides 5ox and 6ox refer to peptides 5 and 6 isolated from CNBr digests of the oxidized protein where their respective levels are enhanced.

20 From the data in Table VII and the comparison of SPG tracks for the oxidized and native protein digests in Figure 8, it is apparent that (1) Met50 is oxidized leading to the loss of peptides X and 9 and the appearance of 5; and (2) Met124 is also oxidized leading to the loss of peptide 7 and the accumulation of peptide 6. Thus oxidation of B. amyloliquifaciens

25 subtilisin with the peracid, diperdocecanoic acid leads to the specific oxidation of methionine at residues 50 and 124.

EXAMPLE 2

30

Substitution at Met50 and Met124  
in Subtilisin Met2220

The choice of amino acid for substitution at Met50 was based on the available sequence data for subtilisins

35

from B. licheniformis (Smith, E.C., et al. (1968) J. Biol. Chem. 243, 2184-2191), B.DY (Nedkov, P., et al. (1983) Hoppe Sayler's Z. Physiol. Chem. 364 1537-1540), B. amylosacchariticus (Markland, F.S., et al. (1967) J. Biol. Chem. 242 5198-5211) and B. subtilis (Stahl, M.L., et al. (1984) J. Bacteriol. 158, 411-418). In all cases, position 50 is a phenylalanine. See Figure 5. Therefore, Phe50 was chosen for construction.

At position 124, all known subtilisins possess a methionine. See Figure 5. Molecular modelling of the x-ray derived protein structure was therefore required to determine the most probable candidates for substitution. From all 19 candidates, isoleucine and leucine were chosen as the best residues to employ. In order to test whether or not modification at one site but not both was sufficient to increase oxidative stability, all possible combinations were built on the Q222 backbone (F50/Q222, I124/Q222, F50/I124/Q222).

#### A. Construction of Mutations Between Codons 45 and 50

All manipulations for cassette mutagenesis were carried out on pS4.5 using methods disclosed in EPO Publication No. 0130756 and Wells, J.A., et al, (1985) Gene 34, 315-323. The pΔ50 in Fig. 10, line 4, mutations was produced using the mutagenesis primer shown in Fig. 10, line 6, and employed an approach designated as restriction-purification which is described below. Briefly, a M13 template containing the subtilisin gene, M13mpl1-SUBT was used for heteroduplex synthesis (Adelman, et al (1983), DNA 2, 183-193). Following transfection of JM101 (ATCC 33876), the 1.5 kb EcoRI-BamHI fragment containing the

subtilisin gene was subcloned from M13mpl1 SUBT rf into a recipient vector fragment of pBS42 the construction of which is described in EPO Publication No. 0130756. To enrich for the mutant sequence (pΔ50, line 4), the resulting plasmid pool was digested with KpnI, and linear molecules were purified by polyacrylamide gel electrophoresis. Linear molecules were ligated back to a circular form, and transformed into E. coli MM294 cells (ATCC 31446). Isolated plasmids were screened by restriction analysis for the KpnI site. KpnI<sup>+</sup> plasmids were sequenced and confirmed the pΔ50 sequence. Asterisks in Figure 11 indicate the bases that are mutated from the wild type sequence (line 4). pΔ50 (line 4) was cut with StuI and EcoRI and the 0.5 Kb fragment containing the 5' half of the subtilisin gene was purified (fragment 1). pΔ50 (line 4) was digested with KpnI and EcoRI and the 4.0 Kb fragment containing the 3' half of the subtilisin gene and vector sequences was purified (fragment 2). Fragments 1 and 2 (line 5), and duplex DNA cassettes coding for mutations desired (shaded sequence, line 6) were mixed in a molar ratio of 1:1:10, respectively. For the particular construction of this example the DNA cassette contained the triplet TTT for codon 50 which encodes Phe. This plasmid was designated pF50. The mutant subtilisin was designated F50.

B. Construction of Mutation  
Between Codons 122 and 127

The procedure of Example 2A was followed in substantial detail except that the mutagenesis primer of Figure 11, line 7 was used and restriction-purification for the EcoRV site in pΔ124 was used. In addition, the DNA cassette (shaded sequence, Figure

11, line 6) contained the triplet ATT for codon 124 which encodes Ile and CTT for Leu. Those plasmids which contained the substitution of Ile for Met124 were designated pI124. The mutant subtilisin was designated I124.

5

C. Construction of Various  
F50/I124/Q222 Multiple Mutants

The triple mutant, F50/I124/Q222, was constructed from a three-way ligation in which each fragment contained one of the three mutations. The single mutant Q222 (pQ222) was prepared by cassette mutagenesis as described in EPO Publication No. 0130756. The F50 mutation was contained on a 2.2kb AvaII to PvuII fragment from pF50; the I124 mutation was contained on a 260 bp PvuII to AvaII fragment from pI124; and the Q222 mutation was contained on 2.7 kb AvaII to AvaII fragment from pQ222. The three fragments were ligated together and transformed into E. coli MM294 cells. Restriction analysis of plasmids from isolated transformants confirmed the construction. To analyze the final construction it was convenient that the AvaII site at position 798 in the wild-type subtilisin gene was eliminated by the I124 construction.

25

The F50/Q222 and I124/Q222 mutants were constructed in a similar manner except that the appropriate fragment from pS4.5 was used for the final construction.

30

D. Oxidative Stability of Q222 Mutants

The above mutants were analyzed for stability to peracid oxidation. As shown in Fig. 12, upon incubation with didecylsuccinic acid (protein 2mg/mL, oxidant 75ppm[O]), both the I124/Q222 and the

35

F50/I124/Q222 are completely stable whereas the F50/Q222 and the Q222 are inactivated. This indicates that conversion of Met124 to I124 in subtilisin Q222 is sufficient to confer resistance to organic peracid oxidants.

5

### EXAMPLE 3

#### 10 Subtilisin Mutants Having Altered Substrate Specificity-Hydrophobic Substitutions at Residues 166

Subtilisin contains an extended binding cleft which is hydrophobic in character. A conserved glycine at residue 166 was replaced with twelve non-ionic amino acids which can project their side-chains into the S-1 subsite. These mutants were constructed to determine the effect of changes in size and hydrophobicity on the binding of various substrates.

20

#### A. Kinetics for Hydrolysis of Substrates Having Altered P-1 Amino Acids by Subtilisin from B. Amyloliquefaciens

Wild-type subtilisin was purified from B. subtilis culture supernatants expressing the B. amylolique-  
25 faciens subtilisin gene (Wells, J.A., et al. (1983) Nucleic Acids Res. 11, 7911-7925) as previously described (Estell, D.A., et al. (1985) J. Biol. Chem. 260, 6518-6521). Details of the synthesis of  
30 tetrapeptide substrates having the form succinyl-L-AlaL-AlaL-ProL-[X]-p-nitroanilide (where X is the P1 amino acid) are described by DelMar, E.G., et al. (1979) Anal. Biochem. 99, 316-320. Kinetic parameters,  $K_m(M)$  and  $k_{cat}(s^{-1})$  were measured using a  
35 modified progress curve analysis (Estell, D.A., et al. (1985) J. Biol. Chem. 260, 6518-6521). Briefly, plots

of rate versus product concentration were fit to the differential form of the rate equation using a non-linear regression algorithm. Errors in  $k_{cat}$  and  $K_m$  for all values reported are less than five percent. The various substrates in Table VIII are ranged in order of decreasing hydrophobicity. Nozaki, Y. (1971), J. Biol. Chem. 246, 2211-2217; Tanford C. (1978) Science 200, 1012).

TABLE VIII

10

	P1 substrate Amino Acid	$k_{cat} (S^{-1})$	$1/K_m (M^{-1})$	$k_{cat}/K_m$ ( $S^{-1}M^{-1}$ )
15	Phe	50	7,100	360,000
	Tyr	28	40,000	1,100,000
	Leu	24	3,100	75,000
	Met	13	9,400	120,000
	His	7.9	1,600	13,000
20	Ala	1.9	5,500	11,000
	Gly	0.003	8,300	21
	Gln	3.2	2,200	7,100
	Ser	2.8	1,500	4,200
	Glu	0.54	32	16

25

The ratio of  $k_{cat}/K_m$  (also referred to as catalytic efficiency) is the apparent second order rate constant for the conversion of free enzyme plus substrate (E+S) to enzyme plus products (E+P) (Jencks, W.P., Catalysis in Chemistry and Enzymology (McGraw-Hill, 1969) pp. 321-436; Fersht, A., Enzyme Structure and Mechanism (Freeman, San Francisco, 1977) pp. 226-287). The log ( $k_{cat}/K_m$ ) is proportional to transition state binding

35

energy,  $\Delta G_T^\ddagger$ . A plot of the  $\log k_{cat}/K_m$  versus the hydrophobicity of the P1 side-chain (Figure 14) shows a strong correlation ( $r = 0.98$ ), with the exception of the glycine substrate which shows evidence for non-productive binding. These data show that relative differences between transition-state binding energies can be accounted for by differences in P-1 side-chain hydrophobicity. When the transition-state binding energies are calculated for these substrates and plotted versus their respective side-chain hydrophobicities, the line slope is 1.2 (not shown). A slope greater than unity, as is also the case for chymotrypsin (Fersht, A., Enzyme Structure and Mechanism (Freeman, San Francisco, 1977) pp. 226-287; Harper, J.W., et al. (1984) Biochemistry, 23, 2995-3002), suggests that the P1 binding cleft is more hydrophobic than ethanol or dioxane solvents that were used to empirically determine the hydrophobicity of amino acids (Nozaki, Y., et al. J. Biol. Chem. (1971) 246, 2211-2217; Tanford, C. (1978) Science 200, 1012).

For amide hydrolysis by subtilisin,  $k_{cat}$  can be interpreted as the acylation rate constant and  $K_m$  as the dissociation constant, for the Michaelis complex (E·S), Ks. Gutfreund, H., et al. (1956) Biochem. J. 63, 656. The fact that the  $\log k_{cat}$ , as well as  $\log 1/K_m$ , correlates with substrate hydrophobicity is consistent with proposals (Robertus, J.D., et al. (1972) Biochemistry 11, 2439-2449; Robertus, J.D., et al. (1972) Biochemistry 11, 4293-4303) that during the acylation step the P-1 side-chain moves deeper into the hydrophobic cleft as the substrate advances from the Michaelis complex (E·S) to the tetrahedral transition-state complex (E·S $^\ddagger$ ). However, these data can also be interpreted as the hydrophobicity of the P1 side-chain effecting the orientation, and thus the

susceptibility of the scissile peptide bond to nucleophilic attack by the hydroxyl group of the catalytic Ser221.

5 The dependence of  $k_{cat}/K_m$  on P-1 side chain hydrophobicity suggested that the  $k_{cat}/K_m$  for hydrophobic substrates may be increased by increasing the hydrophobicity of the S-1 binding subsite. To test this hypothesis, hydrophobic amino acid  
10 substitutions of Gly166 were produced.

Since hydrophobicity of aliphatic side-chains is directly proportional to side-chain surface area (Rose, G.D., et al. (1985) Science 229, 834-838; Reynolds, J.A., et al. (1974) Proc. Natl. Acad. Sci.  
15 USA 71, 2825-2927), increasing the hydrophobicity in the S-1 subsite may also sterically hinder binding of larger substrates. Because of difficulties in predicting the relative importance of these two  
20 opposing effects, we elected to generate twelve non-charged mutations at position 166 to determine the resulting specificities against non-charged substrates of varied size and hydrophobicity.

25 B. Cassette Mutagenesis of the P1 Binding Cleft

The preparation of mutant subtilisins containing the substitution of the hydrophobic amino acids Ala, Val and Phe into residue 166 has been described in EPO  
30 Publication No. 0130756. The same method was used to produce the remaining hydrophobic mutants at residue 166. In applying this method, two unique and silent restriction sites were introduced in the subtilisin genes to closely flank the target codon 166. As can  
35 be seen in Figure 13, the wild type sequence (line 1)



was altered by site-directed mutagenesis in M13 using the indicated 37mer mutagenesis primer, to introduce a 13 bp deletion (dashedline) and unique SacI and XmaI sites (underlined sequences) that closely flank codon 166. The subtilisin gene fragment was subcloned back into the E. coli - B. subtilis shuttle plasmid, pBS42, giving the plasmid p $\Delta$ 166 (Figure 13, line 2). p $\Delta$ 166 was cut open with SacI and XmaI, and gapped linear molecules were purified (Figure 13, line 3). Pools of synthetic oligonucleotides containing the mutation of interest were annealed to give duplex DNA cassettes that were ligated into gapped p $\Delta$ 166 (underlined and overlined sequences in Figure 13, line 4). This construction restored the coding sequence except over position 166(NNN; line 4). Mutant sequences were confirmed by dideoxy sequencing. Asterisks denote sequence changes from the wild type sequence. Plasmids containing each mutant B. amyloliquefaciens subtilisin gene were expressed at roughly equivalent levels in a protease deficient strain of B. subtilis, BG2036 as previously described. EPO Publication No. 0130756; Yang, M., et al. (1984) J. Bacteriol. 160, 15-21; Estell, D.A., et al. (1985) J. Biol. Chem. 260, 6518-6521.

25

#### C. Narrowing Substrate Specificity by Steric Hindrance

To probe the change in substrate specificity caused by steric alterations in the S-1 subsite, position 166 mutants were kinetically analyzed versus P1 substrates of increasing size (i.e., Ala, Met, Phe and Tyr). Ratios of  $k_{cat}/K_m$  are presented in log form in Figure 15 to allow direct comparisons of transition-state binding energies between various enzyme-substrate pairs.

35

According to transition state theory, the free energy difference between the free enzyme plus substrate ( $E + S$ ) and the transition state complex ( $E \cdot S^\ddagger$ ) can be calculated from equation (1),

$$(1) \quad \Delta G_T^\ddagger = -RT \ln k_{cat}/K_m + RT \ln kT/h$$

in which  $k_{cat}$  is the turnover number,  $K_m$  is the Michaelis constant,  $R$  is the gas constant,  $T$  is the temperature,  $k$  is Boltzmann's constant, and  $h$  is Planck's constant. Specificity differences are expressed quantitatively as differences between transition state binding energies (i.e.,  $\Delta\Delta G_T^\ddagger$ ), and can be calculated from equation (2).

$$(2) \quad \Delta\Delta G_T^\ddagger = -RT \ln (k_{cat}/K_m)_A / (k_{cat}/K_m)_B$$

A and B represent either two different substrates assayed against the same enzyme, or two mutant enzymes assayed against the same substrate.

As can be seen from Figure 15A, as the size of the side-chain at position 166 increases the substrate preference shifts from large to small P-1 side-chains. Enlarging the side-chain at position 166 causes  $k_{cat}/K_m$  to decrease in proportion to the size of the P-1 substrate side-chain (e.g., from Gly166 (wild-type) through W166, the  $k_{cat}/K_m$  for the Tyr substrate is decreased most followed in order by the Phe, Met and Ala P-1 substrates).

Specific steric changes in the position 166 side-chain, such as the presence of a  $\beta$ -hydroxyl group,  $\beta$ - or  $\gamma$ -aliphatic branching, cause large decreases in  $k_{cat}/K_m$  for larger P1 substrates. Introducing a  $\beta$ -hydroxyl group in going from A166 (Figure 15A) to

S166 (Figure 15B), causes an 8 fold and 4 fold reduction in  $k_{cat}/K_m$  for Phe and Tyr substrates, respectively, while the values for Ala and Met substrates are unchanged. Producing a  $\beta$ -branched structure, in going from S166 to T166, results in a drop of 14 and 4 fold in  $k_{cat}/K_m$  for Phe and Tyr, respectively. These differences are slightly magnified for V166 which is slightly larger and isosteric with T166. Enlarging the  $\beta$ -branched substituents from V166 to I166 causes a lowering of  $k_{cat}/K_m$  between two and six fold toward Met, Phe and Tyr substrates. Inserting a  $\gamma$ -branched structure, by replacing M166 (Figure 15A) with L166 (Figure 15B), produces a 5 fold and 18 fold decrease in  $k_{cat}/K_m$  for Phe and Tyr substrates, respectively. Aliphatic  $\gamma$ -branched appears to induce less steric hindrance toward the Phe P-1 substrate than  $\beta$ -branching, as evidenced by the 100 fold decrease in  $k_{cat}/K_m$  for the Phe substrate in going from L166 to I166.

Reductions in  $k_{cat}/K_m$  resulting from increases in side chain size in the S-1 subsite, or specific structural features such as  $\beta$ - and  $\gamma$ -branching, are quantitatively illustrated in Figure 16. The  $k_{cat}/K_m$  values for the position 166 mutants determined for the Ala, Met, Phe, and Tyr P-1 substrates (top panel through bottom panel, respectively), are plotted versus the position 166 side-chain volumes (Chothia, C. (1984) Ann. Rev. Biochem. 53, 537-572). Catalytic efficiency for the Ala substrate reaches a maximum for I166, and for the Met substrate it reaches a maximum between V166 and L166. The Phe substrate shows a broad  $k_{cat}/K_m$  peak but is optimal with A166. Here, the  $\beta$ -branched position 166 substitutions form a line that is parallel to, but roughly 50 fold lower in  $k_{cat}/K_m$  than side-chains of similar size [i.e., C166 versus

T166, L166 versus I166]. The Tyr substrate is most efficiently utilized by wild type enzyme (Gly166), and there is a steady decrease as one proceeds to large position 166 side-chains. The  $\beta$ -branched and  $\gamma$ -branched substitutions form a parallel line below the other non-charged substitutions of similar molecular volume.

The optimal substitution at position 166 decreases in volume with increasing volume of the P1 substrate [i.e., I166/Ala substrate, L166/Met substrate, A166/Phe substrate, Gly166/Tyr substrate]. The combined volumes for these optimal pairs may approximate the volume for productive binding in the S-1 subsite. For the optimal pairs, Gly166/Tyr substrate, A166/Phe substrate, L166/Met substrate, V166/Met substrate, and I166/Ala substrate, the combined volumes are 266,295,313,339 and 261  $\text{\AA}^3$ , respectively. Subtracting the volume of the peptide backbone from each pair (i.e., two times the volume of glycine), an average side-chain volume of  $160 \pm 32 \text{\AA}^3$  for productive binding can be calculated.

The effect of volume, in excess to the productive binding volume, on the drop in transition-state binding energy can be estimated from the Tyr substrate curve (bottom panel, Figure 16), because these data, and modeling studies (Figure 2), suggest that any substitution beyond glycine causes steric repulsion. A best-fit line drawn to all the data ( $r = 0.87$ ) gives a slope indicating a loss of roughly 3 kcal/mol in transition state binding energy per  $100 \text{\AA}^3$  of excess volume. ( $100 \text{\AA}^3$  is approximately the size of a leucyl side-chain.)

D. Enhanced Catalytic Efficiency  
Correlates with Increasing Hydrophobicity  
of the Position 166 Substitution

Substantial increases in  $k_{cat}/K_m$  occur with enlargement of the position 166 side-chain, except for the Tyr P-1 substrate (Figure 16). For example,  $k_{cat}/K_m$  increases in progressing from Gly166 to I166 for the Ala substrate (net of ten-fold), from Gly166 to L166 for the Met substrate (net of ten-fold) and from Gly166 to A166 for the Phe substrate (net of two-fold). The increases in  $k_{cat}/K_m$  cannot be entirely explained by the attractive terms in the van der Waals potential energy function because of their strong distance dependence ( $1/r^6$ ) and because of the weak nature of these attractive forces (Jencks, W.P., Catalysis in Chemistry and Enzymology (McGraw-Hill, 1969) pp. 321-436; Fersht, A., Enzyme Structure and Mechanism (Freeman, San Francisco, 1977) pp. 226-287; Levitt, M. (1976) J. Mol. Biol. 104, 59-107). For example, Levitt (Levitt, M. (1976) J. Mol. Biol. 104, 59-107) has calculated that the van der Waals attraction between two methionyl residues would produce a maximal interaction energy of roughly -0.2 kcal/mol. This energy would translate to only 1.4 fold increase in  $k_{cat}/K_m$ .

The increases of catalytic efficiency caused by side-chain substitutions at position 166 are better accounted for by increases in the hydrophobicity of the S-1 subsite. The increase  $k_{cat}/K_m$  observed for the Ala and Met substrates with increasing position 166 side-chain size would be expected, because hydrophobicity is roughly proportional to side-chain surface area (Rose, G.D., et al. (1985) Science 229, 834-838; Reynolds, J.A., et al. (1974) Proc. Natl. Acad. Sci. USA 71, 2825-2927).

Another example that can be interpreted as a hydrophobic effect is seen when comparing  $k_{cat}/K_m$  for isosteric substitutions that differ in hydrophobicity such as S166 and C166 (Figure 16). Cysteine is considerably more hydrophobic than serine (-1.0 versus +0.3 kcal/mol) (Nozaki, Y., et al. (1971) J. Biol. Chem. 246, 2211-2217; Tanford, C. (1978) Science 200, 1012). The difference in hydrophobicity correlates with the observation that C166 becomes more efficient relative to Ser166 as the hydrophobicity of the substrates increases (i.e., Ala < Met < Tye < Phe). Steric hindrance cannot explain these differences because serine is considerably smaller than cysteine (99 versus 118 Å<sup>3</sup>). Paul, I.C., Chemistry of the -SH Group (ed. S. Patai, Wiley Interscience, New York, 1974) pp. 111-149.

E. Production of an Elastase-Like Specificity in Subtilisin

The I166 mutation illustrates particularly well that large changes in specificity can be produced by altering the structure and hydrophobicity of the S-1 subsite by a single mutation (Figure 17). Progressing through the small hydrophobic substrates, a maximal specificity improvement over wild type occurs for the Val substrate (16 fold in  $k_{cat}/K_m$ ). As the substrate side chain size increases, these enhancements shrink to near unity (i.e., Leu and His substrates). The I166 enzyme becomes poorer against larger aromatic substrates of increasing size (e.g., I166 is over 1,000 fold worse against the Tyr substrate than is Gly166). We interpret the increase in catalytic efficiency toward the small hydrophobic substrates for I166 compared to Gly166 to the greater hydrophobicity of isoleucine (i.e., -1.8 kcal/mol versus 0). Nozaki,

Y., et al. (1971) J. Biol. Chem. 246, 2211-2217;  
Tanford, C. (1978) Science 200, 1012. The decrease in  
catalytic efficiency toward the very large substrates  
for I166 versus Gly166 is attributed to steric  
repulsion.

5

The specificity differences between Gly166 and I166  
are similar to the specificity differences between  
chymotrypsin and the evolutionary relative, elastase  
(Harper, J.W., et al (1984) Biochemistry 23,  
10 2995-3002). In elastase, the bulky amino acids, Thr  
and Val, block access to the P-1 binding site for  
large hydrophobic substrates that are preferred by  
chymotrypsin. In addition, the catalytic efficiencies  
toward small hydrophobic substrates are greater for  
15 elastase than for chymotrypsin as we observe for I166  
versus Gly166 in subtilisin.

#### EXAMPLE 4

##### 20 Substitution of Ionic Amino Acids for Gly166

The construction of subtilisin mutants containing the  
substitution of the ionic amino acids Asp, Asn, Gln,  
Lys and Arg are disclosed in EPO Publication No.  
25 0130756. The present example describes the  
construction of the mutant subtilisin containing Glu  
at position 166 (E166) and presents substrate  
specificity data on these mutants. Further data on  
position 166 and 156 single and double mutants is  
30 presented infra.

pA166, described in Example 3, was digested with SacI  
and XmaI. The double strand DNA cassette (underlined  
and overlined) of line 4 in Figure 13 contained the

35

triplet GAA for the codon 166 to encode the replacement of Glu for Gly166. This mutant plasmid designated pQ166 was propagated in BG2036 as described. This mutant subtilisin, together with the other mutants containing ionic substituent amino acids at residue 166, were isolated as described and further analyzed for variations in substrate specificity.

Each of these mutants was analyzed with the tetrapeptide substrates, succinyl-L-AlaL-AlaProL-X-p-nitroanilide, where X was Phe, Ala and Glu.

The results of this analysis are shown in Table IX.

TABLE IX

15

<u>Position 166</u>	<u>P-1 Substrate</u> (kcat/Km x 10 <sup>-4</sup> )		
	<u>Phe</u>	<u>Ala</u>	<u>Glu</u>
20 Gly (wild type)	36.0	1.4	0.002
Asp (D)	0.5	0.4	<0.001
Glu (E)	3.5	0.4	<0.001
Asn (N)	18.0	1.2	0.004
Gln (Q)	57.0	2.6	0.002
25 Lys (K)	52.0	2.8	1.2
Arg (R)	42.0	5.0	0.08

These results indicate that charged amino acid substitutions at Gly166 have improved catalytic efficiencies (kcat/Km) for oppositely charged P-1 substrates (as much as 500 fold) and poorer catalytic efficiency for like charged P-1 substrates.

35



EXAMPLE 5

Substitution of Glycine at Position 169

The substitution of Gly169 in *B. amyloliquefaciens* subtilisin with Ala and Ser is described in EPO  
 5 Publication No. 0130756. The same method was used to make the remaining 17 mutants containing all other substituent amino acids for position 169.

The construction protocol is summarized in Figure 18.  
 10 The overscored and underscored double stranded DNA cassettes used contained the following triplet encoding the substitution of the indicated amino acid at residue 169.

15	GCT	A	ATG	M
	TGT	C	AAC	N
	GAT	D	CCT	P
	GAA	E	CAA	Q
	TTC	F	AGA	R
20	GGC	G	AGC	S
	CAC	H	ACA	T
	ATC	I	GTT	V
	AAA	K	TGG	W
25	CTT	L	TAC	Y

Each of the plasmids containing a substituted Gly169 was designated pX169, where X represents the substituent amino acid. The mutant subtilisins were  
 30 simialrly designated.

Two of the above mutant subtilisins, A169 and S169, were analyzed for substrate specificity against synthetic substrates containing Phe, Leu, Ala and Arg in the P-1 position. The following results are shown  
 35 in Table X.

TABLE X

Effect of Serine and Alanine Mutations  
at Position 169 on P-1 Substrate Specificity

<u>Position 169</u>	<u>P-1 Substrate (kcat/Km x 10<sup>-4</sup>)</u>			
	<u>Phe</u>	<u>Leu</u>	<u>Ala</u>	<u>Arg</u>
Gly (wild type)	40	10	1	0.4
Al69	120	20	1	0.9
S169	50	10	1	0.6

10

These results indicate that substitutions of Ala and Ser at Gly169 have remarkably similar catalytic efficiencies against a range of P-1 substrates compared to their position 166 counterparts. This is probably because position 169 is at the bottom of the P-1 specificity subsite.

EXAMPLE 6

20

Substitution at Position 104

Tyr104 has been substituted with Ala, His, Leu, Met and Ser. The method used was a modification of the site directed mutagenesis method. According to the protocol of Figure 19, a primer (shaded in line 4) introduced a unique HindIII site and a frame shift mutation at codon 104. Restriction-purification for the unique HindIII site facilitated the isolation of the mutant sequence (line 4). Restriction-selection against this HindIII site using primers in line 5 was used to obtain position 104 mutants.

35

The following triplets were used in the primers of Figure 19, line 5 for the 104 codon which substituted the following amino acids.

	GCT	A	TTC	F
5	ATG	M	CCT	P
	CTT	L	ACA	T
	AGC	S	TGG	W
	CAC	H	TAC	Y
	CAA	Q	GTT	V
10	GAA	E	AGA	R
	GGC	G	AAC	N
	ATC	I	GAT	D
	AAA	K	TGT	C

15 The substrates in Table XI were used to analyze the substrate specificity of these mutants. The results obtained fo H104 subtilisin are shown in Table XI.

20 TABLE XI

	<u>Substrate</u>	<u>kcat</u>		<u>Km</u>		<u>Kcat/Km</u>	
		<u>WT</u>	<u>H104</u>	<u>WT</u>	<u>H104</u>	<u>WT</u>	<u>H104</u>
25	sAAPFPNA	50.0	22.0	$1.4 \times 10^{-4}$	$7.1 \times 10^{-4}$	$3.6 \times 10^5$	$3.1 \times 10^4$
	sAAPApNA	3.2	2.0	$2.3 \times 10^{-4}$	$1.9 \times 10^{-3}$	$1.4 \times 10^4$	$1 \times 10^3$
	sFAPFPNA	26.0	38.0	$1.8 \times 10^{-4}$	$4.1 \times 10^{-4}$	$1.5 \times 10^5$	$9.1 \times 10^4$
	sFAPApNA	0.32	2.4	$7.3 \times 10^{-5}$	$1.5 \times 10^{-4}$	$4.4 \times 10^3$	$1.6 \times 10^4$

30 From these data it is clear that the substitution of His for Tyr at position 104 produces an enzyme which is more efficient (higher kcat/Km) when Phe is at the P-4 substrate position than when Ala is at the P-4  
35 substrate position.

EXAMPLE 7Substitution of Ala152

5 Ala152 has been substituted by Gly and Ser to  
determine the effect of such substitutions on  
substrate specificity.

10 The wild type DNA sequence was mutated by the  
V152/P153 primer (Figure 20, line 4) using the above  
restriction-purification approach for the new KpnI  
site. Other mutant primers (shaded sequences Figure  
20; S152, line 5 and G152, line 6) mutated the new  
KpnI site away and such mutants were isolated using  
15 the restriction-selection procedure as described above  
for loss of the KpnI site.

The results of these substitutions for the above  
synthetic substrates containing the P-1 amino acids  
Phe, Leu and Ala are shown in Table XII.  
20

TABLE XII

25 <u>Position 152</u>	<u>P-1 Substrate</u>		
	(kcat/K <sub>m</sub> × 10 <sup>-4</sup> )		
	<u>Phe</u>	<u>Leu</u>	<u>Ala</u>
Gly (G)	0.2	0.4	<0.04
Ala (wild type)	40.0	10.0	1.0
30 Ser (S)	1.0	0.5	0.2

These results indicate that, in contrast to positions  
166 and 169, replacement of Ala152 with Ser or Gly  
35 causes a dramatic reduction in catalytic efficiencies

across all substrates tested. This suggests Ala152, at the top of the S-1 subsite, may be the optimal amino acid because Ser and Gly are homologous Ala substitutes.

5

#### EXAMPLE 8

##### Substitution at Position 156

10 Mutants containing the substitution of Ser and Gln for Glu156 have been constructed according to the overall method depicted in Figure 21. This method was designed to facilitate the construction of multiple mutants at position 156 and 166 as will be described hereinafter. However, by regenerating the wild type  
15 Gly166, single mutations at Glu156 were obtained.

20 The plasmid pA166 is already depicted in line 2 of Figure 13. The synthetic oligonucleotides at the top right of Figure 21 represent the same DNA cassettes depicted in line 4 of Figure 13. The plasmid p166 in Figure 21 thus represents the mutant plasmids of Examples 3 and 4. In this particular example, p166 contains the wild type Gly166.

25 Construction of position 156 single mutants were prepared by ligation of the three fragments (1-3) indicated at the bottom of Figure 21. Fragment 3, containing the carboxy-terminal portion of the subtilisin gene including the wild type position 166  
30 codon, was isolated as a 610 bp SacI-BamHI fragment. Fragment 1 contained the vector sequences, as well as the amino-terminal sequences of the subtilisin gene through codon 151. To produce fragment 1, a unique KpnI site at codon 152 was introduced into the wild  
35 type subtilisin sequence from pS4.5. Site-directed

mutagenesis in M13 employed a primer having the sequence 5'-TA-GTC-GTT-GCG-GTA-CCC-GGT-AAC-GAA-3' to produce the mutation. Enrichment for the mutant sequence was accomplished by restriction with KpnI, purification and self ligation. The mutant sequence containing the KpnI site was confirmed by direct  
5 plasmid sequencing to give pV152. pV152 (~1 µg) was digested with KpnI and treated with 2 units of DNA polymerase I large fragment (Klenow fragment from Boeringer-Mannheim) plus 50 µM deoxynucleotide triphosphates at 37°C for 30 min. This created a  
10 blunt end that terminated with codon 151. The DNA was extracted with 1:1 volumes phenol and CHCl<sub>3</sub> and DNA in the aqueous phase was precipitated by addition of 0.1 volumes 5M ammonium acetate and two volumes ethanol. After centrifugation and washing the DNA pellet with  
15 70% ethanol, the DNA was lyophilized. DNA was digested with BamHI and the 4.6kb piece (fragment 1) was purified by acrylamide gel electrophoresis followed by electroelution. Fragment 2 was a duplex synthetic DNA cassette which when ligated with  
20 fragments 1 and 3 properly restored the coding sequence except at codon 156. The top strand was synthesized to contain a glutamine codon, and the complementary bottom strand coded for serine at 156. Ligation of heterophosphorylated cassettes leads to a  
25 large and favorable bias for the phosphorylated over the non-phosphorylated oligonucleotide sequence in the final segregated plasmid product. Therefore, to obtain Q156 the top strand was phosphorylated, and annealed to the non-phosphorylated bottom strand prior to  
30 ligation. Similarly, to obtain S156 the bottom strand was phosphorylated and annealed to the non-phosphorylated top strand. Mutant sequences were isolated after ligation and transformation, and were confirmed by restriction analysis and DNA sequencing  
35

as before. To express variant subtilisins, plasmids were transformed into a subtilisin-neutral protease deletion mutant of B. subtilis, BG2036, as previously described. Cultures were fermented in shake flasks for 24 h at 37°C in LB media containing 12.5 mg/mL chloramphenicol and subtilisin was purified from  
5 culture supernatants as described. Purity of subtilisin was greater than 95% as judged by SDS PAGE.

These mutant plasmids designated pS156 and pQ156 and mutant subtilisins designated S156 and Q156 were  
10 analyzed with the above synthetic substrates where P-1 comprised the amino acids Glu, Gln, Met and Lys. The results of this analyses are presented in Example 9.

15

#### EXAMPLE 9

##### Multiple Mutants With Altered Substrate Specificity - Substitution at Positions 156 and 166

20

Single substitutions of position 166 are described in Examples 3 and 4. Example 8 describes single substitutions at position 156 as well as the protocol of Figure 21 whereby various double mutants comprising the substitution of various amino acids at positions  
25 156 and 166 can be made. This example describes the construction and substrate specificity of subtilisin containing substitutions at position 156 and 166 and summarizes some of the data for single and double mutants at positions 156 and 166 with various  
30 substrates.

35

K166 is a common replacement amino acid in the 156/166 mutants described herein. The replacement of Lys for

Gly166 was achieved by using the synthetic DNA cassette at the top right of Figure 21 which contained the triplet AAA for NNN. This produced fragment 2 with Lys substituting for Gly166.

5 The 156 substituents were Gln and Ser. The Gln and Ser substitutions at Gly156 are contained within fragment 3 (bottom right Figure 21).

10 The multiple mutants were produced by combining fragments 1, 2 and 3 as described in Example 8. The mutants Q156/K166 and S156/K166 were selectively generated by differential phosphorylation as described. Alternatively, the double 156/166 mutants, c.f. Q156/K166 and S156/K166, were prepared by ligation of the 4.6kb SacI-BamHI fragment from the relevant p156 plasmid containing the 0.6kb SacI-BamHI fragment from the relevant p166 plasmid.

20 These mutants, the single mutant K166, and the S156 and Q156 mutants of Example 8 were analyzed for substitute specificity against synthetic polypeptides containing Phe or Glu as the P-1 substrate residue. The results are presented in Table XIII.

25

30

35



0251446

TABLE XIII

Enzymes Compared (b)	Substrate		Km	kcat/Km	kcat/Km (mutant)	
	P-1 Residue	kcat			kcat/Km (wt)	
Glu156/Gly166 (WT)	Phe	50.00	$1.4 \times 10^{-4}$	$3.6 \times 10^5$	(1)	
	Glu	0.54	$3.4 \times 10^{-2}$	$1.6 \times 10^1$	(1)	
K166	Phe	20.00	$4.0 \times 10^{-5}$	$5.2 \times 10^5$	1.4	
	Glu	0.70	$5.6 \times 10^{-5}$	$1.2 \times 10^4$	750	
Q156/K166	Phe	30.00	$1.9 \times 10^{-5}$	$1.6 \times 10^6$	4.4	
	Glu	1.60	$3.1 \times 10^{-5}$	$5.0 \times 10^4$	3100	
S156/K166	Phe	30.00	$1.8 \times 10^{-5}$	$1.6 \times 10^6$	4.4	
	Glu	0.60	$3.9 \times 10^{-5}$	$1.6 \times 10^4$	1000	
S156	Phe	34.00	$4.7 \times 10^{-5}$	$7.3 \times 10^5$	2.0	
	Glu	0.40	$1.8 \times 10^{-3}$	$1.1 \times 10^2$	6.9	
E156	Phe	48.00	$4.5 \times 10^{-5}$	$1.1 \times 10^6$	3.1	
	Glu	0.90	$3.3 \times 10^{-3}$	$2.7 \times 10^2$	17	

As can be seen in Table XIV, either of these single mutations improve enzyme performance upon substrates with glutamate at the P-1 enzyme binding site. When these single mutations were combined, the resulting multiple enzyme mutants are better than either parent. These single or multiple mutations also alter the relative pH activity profiles of the enzymes as shown in Figure 23.

To isolate the contribution of electrostatics to substrate specificity from other chemical binding forces, these various single and double mutants were analyzed for their ability to bind and cleave synthetic substrates containing Glu, Gln, Met and Lys as the P-1 substrate amino acid. This permitted comparisons between side-chains that were more sterically similar but differed in charge (e.g., Glu versus Gln, Lys versus Met). Similarly, mutant enzymes were assayed against homologous P-1 substrates that were most sterically similar but differed in charge (Table XIV).

TABLE XIV

Kinetics of Position 156/166 Subtilisins  
Determined for Different P1 Substrates

Enzyme Position (a)	Net Charge (b)	P-1 Substrate log kcat/Km (log 1/Km) (c)		
		Glu	Gln	Met
156				
166				
Glu Asp	-2	n.d.	3.02 (2.56)	3.93 (2.74)
Glu Glu	-2	n.d.	3.06 (2.91)	3.86 (3.28)
Glu Asn	-1	1.62 (2.22)	3.85 (3.14)	4.99 (3.85)
Glu Gln	-1	1.20 (2.12)	4.36 (3.64)	5.43 (4.36)
Gln Asp	-1	1.30 (1.79)	3.40 (3.08)	4.94 (3.87)
Ser Asp	-1	1.23 (2.13)	3.41 (3.09)	4.67 (3.68)
Glu Met	-1	1.20 (2.30)	3.89 (3.19)	5.64 (4.83)
Glu Ala	-1	n.d.	4.34 (3.55)	5.65 (4.46)
Glu Gly(wt)	-1	1.20 (1.47)	3.85 (3.35)	5.07 (3.97)
Gln Gly	0	2.42 (2.48)	4.53 (3.81)	5.77 (4.61)
Ser Gly	0	2.31 (2.73)	4.09 (3.68)	5.61 (4.55)
Gln Asn	0	2.04 (2.72)	4.51 (3.76)	5.79 (4.66)
Ser Asn	0	1.91 (2.78)	4.57 (3.82)	5.72 (4.64)
Glu Arg	0	2.91 (3.30)	4.26 (3.50)	5.32 (4.22)
Glu Lys	0	4.09 (4.25)	4.70 (3.88)	6.15 (4.45)
Gln Lys	+1	4.70 (4.50)	4.64 (3.68)	5.97 (4.68)
Ser Lys	+1	4.21 (4.40)	4.84 (3.94)	6.16 (4.90)
Maximum difference:				
log kcat/Km (log 1/Km) (d)		3.5 (3.0)	1.8 (1.4)	2.3 (2.2)
				-1.3 (-1.0)

Footnotes to Table XIV:

- (a) B. subtilis, BG 2036, expressing indicated variant subtilisin were fermented and enzymes purified as previously described (Estell, et al. (1985) J. Biol. Chem. 260, 6518-6521). Wild type subtilisin is indicated (wt) containing Glu156 and Gly166.
- (b) Net charge in the P-1 binding site is defined as the sum of charges from positions 156 and 166 at pH 8.6.
- (c) Values for  $k_{cat}(s^{-1})$  and  $K_m(M)$  were measured in 0.1M Tris pH 8.6 at 25°C as previously described against P-1 substrates having the form succinyl-L-AlaL-AlaL-ProL-[X]-p-nitroanilide, where X is the indicated P-1 amino acid. Values for  $\log 1/K_m$  are shown inside parentheses. All errors in determination of  $k_{cat}/K_m$  and  $1/K_m$  are below 5%.
- (d) Because values for Glu156/Asp166(D166) are too small to determine accurately, the maximum difference taken for GluP-1 substrate is limited to a charge range of +1 to -1 charge change.
- n.d. = not determined

The  $k_{cat}/K_m$  ratios shown are the second order rate constants for the conversion of substrate to product, and represent the catalytic efficiency of the enzyme. These ratios are presented in logarithmic form to scale the data, and because  $\log k_{cat}/K_m$  is proportional to the lowering of transition-state activation energy ( $\Delta G_T$ ). Mutations at position 156 and 166 produce changes in catalytic efficiency toward Glu, Gln, Met and Lys P-1 substrates of 3100, 60, 200 and 20 fold, respectively. Making the P-1 binding-site more positively charged [e.g., compare Gln156/Lys166 (Q156/K166) versus Glu156/Met166 (Glu156/M166)] dramatically increased  $k_{cat}/K_m$  toward the Glu P-1 substrate (up to 3100 fold), and decreased the catalytic efficiency toward the Lys P-1 substrate (up to 10 fold). In addition, the results show that the catalytic efficiency of wild type enzyme can be

greatly improved toward any of the four P-1 substrates by mutagenesis of the P-1 binding site.

5 The changes in  $k_{cat}/K_m$  are caused predominantly by changes in  $1/K_m$ . Because  $1/K_m$  is approximately equal to  $1/K_s$ , the enzyme-substrate association constant, the mutations primarily cause a change in substrate binding. These mutations produce smaller effects on  $k_{cat}$  that run parallel to the effects on  $1/K_m$ . The changes in  $k_{cat}$  suggest either an alteration in binding in the P-1 binding site in going from the Michaelis-complex E·S to the transition-state complex (E-S<sup>‡</sup>) as previously proposed (Robertus, J.D., et al. (1972) Biochemistry 11, 2439-2449; Robertus, J.D., et al. (1972) Biochemistry 11, 4293-4303), or change in the position of the scissile peptide bond over the catalytic serine in the E·S complex.

20 Changes in substrate preference that arise from changes in the net charge in the P-1 binding site show trends that are best accounted for by electrostatic effects (Figure 28). As the P-1 binding cleft becomes more positively charged, the average catalytic efficiency increases much more for the Glu P-1 substrate than for its neutral and isosteric P-1 homolog, Gln (Figure 28A). Furthermore, at the positive extreme both substrates have nearly identical catalytic efficiencies.

30 In contrast, as the P-1 site becomes more positively charged the catalytic efficiency toward the Lys P-1 substrate decreases, and diverges sharply from its neutral and isosteric homolog, Met (Figure 28B). The similar and parallel upward trend seen with increasing positive charge for the Met and Glu P-1 substrates probably results from the fact that all the substrates

are succinylated on their amino-terminal end, and thus carry a formal negative charge.

The trends observed in  $\log k_{cat}/K_m$  are dominated by changes in the  $K_m$  term (Figures 28C and 28D). As the pocket becomes more positively charged, the  $\log 1/K_m$  values converge for Glu and Gln P-1 substrates (Figure 28C), and diverge for Lys and Met P-1 substrates (Figure 28D). Although less pronounced effects are seen in  $\log k_{cat}$ , the effects of P-1 charge on  $\log k_{cat}$  parallel those seen in  $\log 1/K_m$  and become larger as the P-1 pocket becomes more positively charged. This may result from the fact that the transition-state is a tetrahedral anion, and a net positive charge in the enzyme may serve to provide some added stabilization to the transition-state.

The effect of the change in P-1 binding-site charge on substrate preference can be estimated from the differences in slopes between the charged and neutral isosteric P-1 substrates (Figure 28B). The average change in substrate preference ( $\Delta \log k_{cat}/K_m$ ) between charged and neutral isosteric substrates increases roughly 10-fold as the complementary charge or the enzyme increases (Table XV). When comparing Glu versus Lys, this difference is 100-fold and the change in substrate preference appears predominantly in the  $K_m$  term.

30

35

TABLE XV

0251446

Differential Effect on Binding Site  
Charge on log kcat/Km or (log 1/Km) (a)  
for P-1 Substrates that Differ in Charge

5	Change in P-1 Binding Site Charge (b)	$\Delta \log$ kcat/Km		( $\Delta \log$ 1/Km)
		GluGln	MetLys	GluLys
	-2 to -1	n.d.	1.2 (1.2)	n.d.
	-1 to 0	0.7 (0.6)	1.3 (0.8)	2.1 (1.4)
	0 to +1	1.5 (1.3)	0.5 (0.3)	2.0 (1.5)
10	Avg. change in log kcat/K <sub>m</sub> or (log 1/Km) <sub>m</sub> per unit charge change			
		1.1 (1.0)	1.0 (0.8)	2.1 (1.5)

15 (a) The difference in the slopes of curves were taken between the P-1 substrates over the charge interval given for log (kcat/Km) (Figure 28A, B) and (log 1/Km) (Figure 28C, D). Values represent the differential effect a charge change has in distinguishing the substrates that are compared.

20 (b) Charge in P-1 binding site is defined as the sum of charges from positions 156 and 166.

25

30

35

The free energy of electrostatic interactions in the structure and energetics of salt-bridge formation depends on the distance between the charges and the microscopic dielectric of the media. To dissect these structural and microenvironmental effects, the energies involved in specific salt-bridges were evaluated. In addition to the possible salt-bridges shown (Figures 29A and 29B), reasonable salt-bridges can be built between a Lys P-1 substrate and Asp at position 166, and between a Glu P-1 substrate and a Lys at position 166 (not shown). Although only one of these structures is confirmed by X-ray crystallography (Poulos, T.L., et al. (1976) J. Mol. Biol. 257 1097-1103), all models have favorable torsion angles (Sielecki, A.R., et al. (1979) J. Mol. Biol. 134, 781-804), and do not introduce unfavorable van der Waals contacts.

The change in charged P-1 substrate preference brought about by formation of the model salt-bridges above are shown in Table XVI.



TABLE XVI

Effect of Salt Bridge Formation Between Enzyme  
and Substrate on P1 Substrate Preference (a)

Enzymes Compared (b)		Enzyme Position Changed	P-1 Substrates Compared	Substrate Preference $\Delta \log$ (kcat/Km)		Change in Substrate Preference $\Delta \Delta \log$ (kcat/Km)
1	2			1	2	
Glu156/Asp166	Gln156/Asp166	156	LysMet	+0.30	-0.53	0.83
Glu156/Asn166	Gln156/Asn166	156	LysMet	-0.84	-2.04	1.20
Glu156/Gly166	Gln156/Gly166	156	LysMet	-0.47	-2.10	1.63
Glu156/Lys-166	Gln156/Lys166	156	LysMet	-1.92	-2.74	0.82
				Ave $\Delta \Delta \log$ (kcat/Km)		1.10 $\pm$ 0.3
Glu156/Asp166	Glu156/Asn166	166	LysMet	+0.30	-0.84	1.14
Glu156/Glu166	Glu156/Glu166	166	LysMet	+0.62	-1.33	1.95
Gln156/Asp166	Gln156/Asn166	166	LysMet	-0.53	-2.04	1.51
Ser156/Asp166	Ser156/Asn166	166	LysMet	-0.43	-2.04	1.61
Glu156/Lys166	Glu156/Met166	166	GluGln	-0.63	-2.69	2/06
				Ave $\Delta \Delta \log$ (kcat/Km)		1.70 $\pm$ 0.3

Footnotes to Table XVI:

- (a) Molecular modeling shows it is possible to form a salt bridge between the indicated charged P-1 substrate and a complementary charge in the P-1 binding site of the enzyme at the indicated position changed.
- 5 (b) Enzymes compared have sterically similar amino acid substitutions that differ in charge at the indicated position.
- (c) The P-1 substrates compared are structurally similar but differ in charge. The charged P-1 substrate is complementary to the charge change at the position indicated between enzymes 1 and 2.
- 10 (d) Data from Table XIV was used to compute the difference in  $\log(k_{cat}/K_m)$  between the charged and the non-charged P-1 substrate (i.e., the substrate preference). The substrate preference is shown separately for enzyme 1 and 2.
- 15 (e) The difference in substrate preference between enzyme 1 (more highly charged) and enzyme 2 (more neutral) represents the rate change accompanying the electrostatic interaction.
- 
- 20 The difference between catalytic efficiencies (i.e.,  $\Delta \log k_{cat}/K_m$ ) for the charged and neutral P-1 substrates (e.g., Lys minus Met or Glu minus Gln) give the substrate preference for each enzyme. The change in substrate preference ( $\Delta \Delta \log k_{cat}/K_m$ ) between the charged and more neutral enzyme homologs (e.g.,
- 25 Glu156/Gly166 minus Gln156(Q156)/Gly166) reflects the change in catalytic efficiency that may be attributed solely to electrostatic effects.
- 30 These results show that the average change in substrate preference is considerably greater when electrostatic substitutions are produced at position 166 (50-fold in  $k_{cat}/K_m$ ) versus position 156 (12-fold in  $k_{cat}/K_m$ ). From these  $\Delta \log k_{cat}/K_m$  values, an average change in transition-state stabilization
- 35 energy can be calculated of -1.5 and -2.4 kcal/mol for

substitutions at positions 156 and 166, respectively. This should represent the stabilization energy contributed from a favorable electrostatic interaction for the binding of free enzyme and substrate to form the transition-state complex.

5

EXAMPLE 10

Substitutions at Position 217

10 Tyr217 has been substituted by all other 19 amino acids. Cassette mutagenesis as described in EPO publication No. 0130756 was used according to the protocol of Figure 22. The EcoRV restriction site was used for restriction-purification of pΔ217.

15 Since this position is involved in substrate binding, mutations here effect kinetic parameters of the enzyme. An example is the substitution of Leu for Tyr at position 217. For the substrate sAAPFpNa, this mutant has a kcat of 277 5' and a Km of  $4.7 \times 10^{-4}$  with  
20 a kcat/Km ratio of  $6 \times 10^5$ . This represents a 5.5-fold increase in kcat with a 3-fold increase in Km over the wild type enzyme.

25 In addition, replacement of Tyr217 by Lys, Arg, Phe or Leu results in mutant enzymes which are more stable at pHs of about 9-11 than the WT enzyme. Conversely, replacement of Tyr217 by Asp, Glu, Gly or Pro results in enzymes which are less stable at pHs of about 9-11  
30 than the WT enzyme.

35

EXAMPLE 11

Multiple Mutants Having  
Altered Thermal Stability

5 B. amyloliquefacien subtilisin does not contain any  
cysteine residues. Thus, any attempt to produce  
thermal stability by Cys cross-linkage required the  
substitution of more than one amino acid in subtilisin  
with Cys. The following subtilisin residues were  
multiply substituted with cysteine:

10 Thr22/Ser87  
Ser24/Ser87

15 Mutagenesis of Ser24 to Cys was carried out with a 5'  
phosphorylated oligonucleotide primer having the  
sequence

5'-pC-TAC-ACT-GGA-TGC<sup>\*\*</sup>-AAT-GTT-AAA-G-3'.

20 (Asterisks show the location of mismatches and the  
underlined sequence shows the position of the altered  
Sau3A site.) The B. amyloliquefaciens subtilisin gene  
on a 1.5 kb EcoRI-BAMHI fragment from pS4.5 was cloned  
into M13mp11 and single stranded DNA was isolated.  
This template (M13mp11SUBT) was double primed with the  
25 5' phosphorylated M13 universal sequencing primer and  
the mutagenesis primer. Adelman, et al. (1983) DNA 2,  
183-193. The heteroduplex was transfected into  
competent JM101 cells and plaques were probed for the  
mutant sequence (Zoller, M.J., et al. (1982) Nucleic  
30 Acid Res. 10, 6487-6500; Wallace, et al. (1981)  
Nucleic Acid Res. 9, 3647-3656) using a  
tetramethylammonium chloride hybridization protocol  
(Wood, et al. (1985) Proc. Natl. Acad. Sci. USA 82,  
1585-1588). The Ser87 to Cys mutation was prepared in

a similar fashion using a 5' phosphorylated primer having the sequence

5'-pGGC-GTT-GCG-CCA-<sup>\*</sup>TGC-GCA-TCA-CT-3'.

5 (The asterisk indicates the position of the mismatch and the underlined sequence shows the position of a new MstI site.) The C24 and C87 mutations were obtained at a frequency of one and two percent, respectively. Mutant sequences were confirmed by dideoxy sequencing in M13.

10 Mutagenesis of Tyr21/Thr22 to A21/C22 was carried out with a 5' phosphorylated oligonucleotide primer having the sequence

15 5'-pAC-TCT-CAA-GGC-<sup>\*\*\*</sup>GCT-TGT-<sup>\*\*</sup>GGC-TCA-AAT-GTT-3'.

20 (The asterisks show mismatches to the wild type sequence and the underlined sequence shows the position of an altered Sau3A site.) Manipulations for heteroduplex synthesis were identical to those described for C24. Because direct cloning of the heteroduplex DNA fragment can yield increased frequencies of mutagenesis, the EcoRI-BamHI subtilisin fragment was purified and ligated into pBS42. E. coli MM 294 cells were transformed with the ligation mixture and plasmid DNA was purified from isolated transformants. Plasmid DNA was screened for the loss of the Sau3A site at codon 23 that was eliminated by the mutagenesis primer. Two out of 16 plasmid preparations had lost the wild type Sau3A site. The mutant sequence was confirmed by dideoxy sequencing in M13.

Double mutants, C22/C87 and C24/C87, were constructed by ligating fragments sharing a common ClaI site that separated the single parent cystine codons. Specifically, the 500 bp EcoRI-ClaI fragment containing the 5' portion of the subtilisin gene (including codons 22 and 24) was ligated with the 4.7 kb ClaI-EcoRI fragment that contained the 3' portion of the subtilisin gene (including codon 87) plus pBS42 vector sequence. E. coli MM 294 was transformed with ligation mixtures and plasmid DNA was purified from individual transformants. Double-cysteine plasmid constructions were identified by restriction site markers originating from the parent cysteine mutants (i.e., C22 and C24, Sau3A minus; Cys87, MstI plus). Plasmids from E. coli were transformed into B. subtilis BG2036. The thermal stability of these mutants as compared to wild type subtilisin are presented in Figure 30 and Tables XVII and XVIII.

TABLE XVII

Effect of DTT on the Half-Time of  
Autolytic Inactivation of Wild-Type  
and Disulfide Mutants of Subtilisin\*

Enzyme	$t_{1/2}$		-DTT/+DTT
	-DDT	+DTT	
	min		
Wild-type	95	85	1.1
C22/C87	44	25	1.8
C24/C87	92	62	1.5

(\*) Purified enzymes were either treated or not treated with 25mM DTT and dialyzed with or without 10mM DTT in 2mM  $\text{CaCl}_2$ , 50mM Tris (pH 7.5) for 14 hr. at 4°C. Enzyme concentrations were adjusted to 80 $\mu$ l aliquots were quenched on ice and assayed for residual activity. Half-times for autolytic inactivation were determined from semi-log plots of  $\log_{10}$  (residual activity) versus time. These plots were linear for over 90% of the inactivation.

TABLE XVIII

Effect of Mutations in Subtilisin  
on the Half-Time of Autolytic  
Inactivation at 58°C\*

5	Enzyme	$t_{1/2}$
		min
	Wild-type	120
	C22	22
	C24	120
	C87	104
10	C22/C87	43
	C24/C87	115

(\*) Half-times for autolytic inactivation were determined for wild-type and mutant subtilisins as described in the legend to Table III. Unpurified and non-reduced enzymes were used directly from B. subtilis culture supernatants.

The disulfides introduced into subtilisin did not improve the autolytic stability of the mutant enzymes when compared to the wild-type enzyme. However, the disulfide bonds did provide a margin of autolytic stability when compared to their corresponding reduced double-cysteine enzyme. Inspection of a highly refined x-ray structure of wild-type B. amyloliquefaciens subtilisin reveals a hydrogen bond between Thr22 and Ser87. Because cysteine is a poor hydrogen donor or acceptor (Paul, I.C. (1974) in Chemistry of the -SH Group (Patai, S., ed.) pp. 111-149, Wiley Interscience, New York) weakening of 22/87 hydrogen bond may explain why the C22 and C87 single-cysteine mutant proteins are less autolytically stable than either C24 or wild-type (Table XVIII). The fact that C22 is less autolytically stable than C87 may be the result of the Tyr21A mutation (Table XVIII). Indeed,



construction and analysis of Tyr21/C22 shows the mutant protein has an autolytic stability closer to that of C87. In summary, the C22 and C87 of single-cysteine mutations destabilize the protein toward autolysis, and disulfide bond formation  
5 increases the stability to a level less than or equal to that of wild-type enzyme.

#### EXAMPLE 12

##### 10 Multiple Mutants Containing Substitutions at Position 222 and Position 166 or 169

Double mutants 166/222 and 169/222 were prepared by ligating together (1) the 2.3kb AcaII fragment from  
15 pS4.5 which contains the 5' portion of the subtilisin gene and vector sequences, (2) the 200bp AvaII fragment which contains the relevant 166 or 169 mutations from the respective 166 or 169 plasmids, and  
20 (3) the 2.2kb AvaII fragment which contains the relevant 222 mutation 3' and of the subtilisin genes and vector sequence from the respective p222 plasmid.

Although mutations at position 222 improve oxidation stability they also tend to increase the  $K_m$ . An  
25 example is shown in Table XIX. In this case the A222 mutation was combined with the K166 mutation to give an enzyme with  $k_{cat}$  and  $K_m$  intermediate between the two parent enzymes.

30

35

TABLE XIX

	<u>kcat</u>	<u>Km</u>
WT	50	$1.4 \times 10^{-4}$
A222	42	$9.9 \times 10^{-4}$
5 K166	21	$3.7 \times 10^{-5}$
K166/A222	29	$2.0 \times 10^{-4}$

substrate sAAPFpNa

10

EXAMPLE 13

15 Multiple Mutants Containing  
Substitutions at Positions 50, 156,  
166, 217 and Combinations Thereof

20 The double mutant S156/A169 was prepared by ligation  
of two fragments, each containing one of the relevant  
mutations. The plasmid pS156 was cut with XmaI and  
25 treated with S1 nuclease to create a blunt end at  
codon 167. After removal of the nuclease by  
phenol/chloroform extraction and ethanol precipita-  
tion, the DNA was digested with BamHI and the  
approximately 4kb fragment containing the vector plus  
the 5' portion of the subtilisin gene through codon  
167 was purified.

30 The pA169 plasmid was digested with KpnI and treated  
with DNA polymerase Klenow fragment plus 50  $\mu$ M dNTPs  
to create a blunt end codon at codon 168. The Klenow  
was removed by phenol/chloroform extraction and  
ethanol precipitation. The DNA was digested with  
BamHI and the 590bp fragment including codon 168  
through the carboxy terminus of the subtilisin gene

35

was isolated. The two fragments were then ligated to give S156/A169.

Triple and quadruple mutants were prepared by ligating together (1) the 220bp PvuII/HaeII fragment containing the relevant 156, 166 and/or 169 mutations from the  
5 respective p156, p166 and/or p169 double of single mutant plasmid, (2) the 550bp HaeII/BamHI fragment containing the relevant 217 mutant from the respective p217 plasmid, and (3) the 3.9kb PvuII/BamHI fragment containing the F50 mutation and vector sequences.

10 The multiple mutant F50/S156/A169/L217, as well as B. amyloliquefaciens subtilisin, B. lichenformis subtilisin and the single mutant L217 were analyzed with the above synthetic polypeptides where the P-1  
15 amino acid in the substrate was Lys, His, Ala, Gln, Tyr, Phe, Met and Leu. These results are shown in Figures 26 and 27.

20 These results show that the F50/S156/A169/L217 mutant has substrate specificity similar to that of the B. licheniformis enzyme and differs dramatically from the wild type enzyme. Although only data for the L217 mutant are shown, none of the single mutants (e.g., F50, S156 or A169) showed this effect. Although B.  
25 licheniformis differs in 88 residue positions from B. amyloliquefaciens, the combination of only these four mutations accounts for most of the differences in substrate specificity between the two enzymes.

#### 30 EXAMPLE 14

##### Subtilisin Mutants Having Altered Alkaline Stability

A random mutagenesis technique was used to generate single and multiple mutations within the B.  
35

amyloliquefaciens subtilisin gene. Such mutants were screened for altered alkaline stability. Clones having increased (positive) alkaline stability and decreased (negative) alkaline stability were isolated and sequenced to identify the mutations within the subtilisin gene. Among the positive clones, the mutants V107 and R213 were identified. These single mutants were subsequently combined to produce the mutant V107/R213.

One of the negative clones (V50) from the random mutagenesis experiments resulted in a marked decrease in alkaline stability. Another mutant (P50) was analyzed for alkaline stability to determine the effect of a different substitution at position 50. The F50 mutant was found to have a greater alkaline stability than wild type subtilisin and when combined with the double mutant V107/R213 resulted in a mutant having an alkaline stability which reflected the aggregate of the alkaline stabilities for each of the individual mutants.

The single mutant R204 and double mutant C204/R213 were identified by alkaline screening after random cassette mutagenesis over the region from position 197 to 228. The C204/R213 mutant was thereafter modified to produce mutants containing the individual mutations C204 and R213 to determine the contribution of each of the individual mutations. Cassette mutagenesis using pooled oligonucleotides to substitute all amino acids at position 204, was utilized to determine which substitution at position 204 would maximize the increase in alkaline stability. The mutation from Lys213 to Arg was maintained constant for each of these substitutions at position 204.

A. Construction of pB0180, an  
E. coli-B. subtilis Shuttle Plasmid

The 2.9 kb EcoRI-BamHI fragment from pBR327 (Covarrubias, L., et al. (1981) Gene 13, 25-35) was ligated to the 3.7kb EcoRI-BamHI fragment of pBD64 (Gryczan, T., et al. (1980) J. Bacteriol., 141, 246-253) to give the recombinant plasmid pB0153. The unique EcoRI recognition sequence in pBD64 was eliminated by digestion with EcoRI followed by treatment with Klenow and deoxynucleotide triphosphates (Maniatis, T., et al. (eds.) (1982) in Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.). Blunt end ligation and transformation yielded pB0154. The unique AvaI recognition sequence in pB0154 was eliminated in a similar manner to yield pB0171. pB0171 was digested with BamHI and PvuII and treated with Klenow and deoxynucleotide triphosphates to create blunt ends. The 6.4 kb fragment was purified, ligated and transformed into LE392 cells (Enquest, L.W., et al. (1977) J. Mol. Biol. 111, 97-120), to yield pB0172 which retains the unique BamHI site. To facilitate subcloning of subtilisin mutants, a unique and silent KpnI site starting at codon 166 was introduced into the subtilisin gene from pS4.5 (Wells, J.A., et al. (1983) Nucleic Acids Res., 11, 7911-7925) by site-directed mutagenesis. The KpnI+ plasmid was digested with EcoRI and treated with Klenow and deoxynucleotide triphosphates to create a blunt end. The Klenow was inactivated by heating for 20 min at 68°C, and the DNA was digested with BamHI. The 1.5 kb blunt EcoRI-BamHI fragment containing the entire subtilisin was ligated with the 5.8 kb NruI-BamHI from pB0172 to yield pB0180. The ligation of the blunt NruI end to the blunt EcoRI end recreated an EcoRI

5 site. Proceeding clockwise around pB0180 from the EcoRI site at the 5' end of the subtilisin gene is the unique BamHI site at the 3' end of the subtilisin gene, the chloramphenicol and neomycin resistance genes and UB110 gram positive replication origin derived from pBD64, the ampicillin resistance gene and gram negative replication origin derived from pBR327.

#### B. Construction of Random Mutagenesis Library

10 The 1.5 kb EcoRI-BamHI fragment containing the B. amyloliquefaciens subtilisin gene (Wells et al., 1983) from pB0180 was cloned into M13mp11 to give M13mp11 SUBT essentially as previously described (Wells, J.A.,  
15 et al. (1986) J. Biol. Chem., 261,6564-6570). Deoxyuridine containing template DNA was prepared according to Kunkel (Kunkel, T.A. (1985) Proc. Natl. Acad. Sci. USA, 82 488-492). Uridine containing  
20 template DNA (Kunkel, 1985) was purified by CsCl density gradients (Maniatis, T. et al. (eds.) (1982) in Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.). A primer (AvaI<sup>-</sup>) having the sequence

25 5'GAAAAAAGACCCTAGCGTCGCTTA

ending at codon -11, was used to alter the unique AvaI recognition sequence within the subtilisin gene. (The asterisk denotes the mismatches from the wild-type sequence and underlined is the altered AvaI site.)  
30

The 5' phosphorylated AvaI primer (~320 pmol) and ~40 pmol (~120µg) of uridine containing M13mp11 SUBT template in 1.88 ml of 53 mM NaCl, 7.4 mM MgCl<sub>2</sub> and 7.4 mM Tris.HCl (pH 7.5) were annealed by heating to  
35

90°C for 2 min. and cooling 15 min at 24°C (Fig. 31). Primer extension at 24°C was initiated by addition of 100µL containing 1 mM in all four deoxynucleotide triphosphates, and 20µl Klenow fragment (5 units/l). The extension reaction was stopped every 15 seconds over ten min by addition of 10µl 0.25 M EDTA (pH 8) to 50µl aliquots of the reaction mixture. Samples were pooled, phenol chlorophorm extracted and DNA was precipitated twice by addition of 2.5 vol 100% ethanol, and washed twice with 70% ethanol. The pellet was dried, and redissolved in 0.4 ml 1 mM EDTA, 10 mM Tris (pH 8).

Misincorporation of α-thiodeoxynucleotides onto the 3' ends of the pool of randomly terminated template was carried out by incubating four 0.2 ml solutions each containing one-fourth of the randomly terminated template mixture (~20µg), 0.25 mM of a given α-thiodeoxynucleotide triphosphate, 100 units AMV polymerase, 50 mM KCL, 10 mM MgCl<sub>2</sub>, 0.4 mM dithiothreitol, and 50 mM Tris (pH 8.3) (Champoux, J.J. (1984) Genetics, 2, 454-464). After incubation at 37°C for 90 minutes, misincorporation reactions were sealed by incubation for five minutes at 37°C with 50 mM all four deoxynucleotide triphosphates (pH 8), and 50 units AMV polymerase. Reactions were stopped by addition of 25 mM EDTA (final), and heated at 68°C for ten min to inactivate AMV polymerase. After ethanol precipitation and resuspension, synthesis of closed circular heteroduplexes was carried out for two days at 14°C under the same conditions used for the timed extension reactions above, except the reactions also contained 1000 units T4 DNA ligase, 0.5 mM ATP and 1 mM β-mercaptoethanol. Simultaneous restriction of each heteroduplex pool with KpnI, BamHI, and EcoRI confirmed that the

extension reactions were nearly quantitative. Heteroduplex DNA in each reaction mixture was methylated by incubation with 80 $\mu$ M S-adenosylmethionine and 150 units dam methylase for 1 hour at 37°C. Methylation reactions were stopped by heating at 68°C for 15 min.

5

One-half of each of the four methylated heteroduplex reactions were transformed into 2.5 ml competent E. coli JM101 (Messing, J. (1979) Recombinant DNA Tech. Bull., 2, 43-48). The number of independent transformants from each of the four transformations ranged from 0.4-2.0 x 10<sup>5</sup>. After growing out phage pools, RF DNA from each of the four transformations was isolated and purified by centrifugation through CsCl density gradients. Approximately 2 $\mu$ g of RF DNA from each of the four pools was digested with EcoRI, BamHI and AvaI. The 1.5 kb EcoRI-BamHI fragment (i.e., AvaI resistant) was purified on low gel temperature agarose and ligated into the 5.5 kb EcoRI-BamHI vector fragment of pB0180. The total number of independent transformants from each  $\alpha$ -thiodeoxynucleotide misincorporation plasmid library ranged from 1.2-2.4 x 10<sup>4</sup>. The pool of plasmids from each of the four transformations was grown out in 200 ml LB media containing 12.5 $\mu$ g/ml cmp and plasmid DNA was purified by centrifugation through CsCl density gradients.

15

20

25

### 30 C. Expression and Screening of Subtilisin Point Mutants

Plasmid DNA from each of the four misincorporation pools was transformed (Anagnostopoulos, C., et al. (1967), J. Bacteriol., 81, 741-746) into BG2036. For each transformation, 5 $\mu$ g of DNA produced approximately

35



2.5 x 10<sup>5</sup> independent BG2036 transformants, and liquid culture aliquots from the four libraries were stored in 10% glycerol at 70°C. Thawed aliquots of frozen cultures were plated on LB/5µg/ml cmp/1.6% skim milk plates (Wells, J.A., et al. (1983) Nucleic Acids Res., 11, 7911-7925), and fresh colonies were arrayed onto 5 96-well microtiter plates containing 150 µl per well LB media plus 12.5µg/ml cmp. After 1 h at room temperature, a replica was stamped (using a matched 96 prong stamp) onto a 132 mm BA 85 nitrocellulose filter (Schleicher and Scheull) which was layered on a 140 mm 10 diameter LB/cmp/skim milk plate. Cells were grown about 16 h at 30°C until halos of proteolysis were roughly 5-7 mm in diameter and filters were transferred directly to a freshly prepared agar plate at 37°C containing only 1.6% skim milk and 50 mM 15 sodium phosphate pH 11.5. Filters were incubated on plates for 3-6 h at 37°C to produce halos of about 5 mm for wild-type subtilisin and were discarded. The plates were stained for 10 min at 24°C with Coomassie blue solution (0.25% Coomassie blue (R-250) 25% 20 ethanol) and destained with 25% ethanol, 10% acetic acid for 20 min. Zones of proteolysis appeared as blue halos on a white background on the underside of the plate and were compared to the original growth plate that was similarly stained and destained as a 25 control. Clones were considered positive that produced proportionately larger zones of proteolysis on the high pH plates relative to the original growth plate. Negative clones gave smaller halos under alkaline conditions. Positive and negative clones 30 were restreaked to colony purify and screened again in triplicate to confirm alkaline pH results.

D. Identification and Analysis  
of Mutant Subtilisins

Plasmid DNA from 5 ml overnight cultures of more alkaline active B.subtilis clones was prepared according to Birnboim and Doly (Birnboim, H.C., et al. (1979) Nucleic Acid Res. 7, 1513) except that

5 incubation with 2 mg/ml lysozyme proceeded for 5 min at 37°C to ensure cell lysis and an additional phenol/CHCl<sub>3</sub> extraction was employed to remove contaminants. The 1.5 kb EcoRI-BamHI fragment containing the subtilisin gene was ligated into

10 M13mp11 and template DNA was prepared for DNA sequencing (Messing, J., et al. (1982) Gene, 19 269-276). Three DNA sequencing primers ending at codon 26, +95, and +155 were synthesized to match the subtilisin coding sequence. For preliminary sequence

15 identification a single track of DNA sequence, corresponding to the dNTPas misincorporation library from which the mutant came, was applied over the entire mature protein coding sequence (i.e., a single dideoxyguanosine sequence track was applied to

20 identify a mutant from the dGTPas library). A complete four track of DNA sequence was performed 200 bp over the site of mutagenesis to confirm and identify the mutant sequence (Sanger, F., et al., (1980) J. Mol. Biol., 143, 161-178). Confirmed

25 positive and negative bacilli clones were cultured in LB media containing 12.5µg/mL cmp and purified from culture supernatants as previously described (Estell, D.A., et al. (1985) J. Biol. Chem., 260, 6518-6521). Enzymes were greater than 98% pure as analyzed by

30 SDS-polyacrylamide gel electrophoresis (Laemmli, U.K. (1970), Nature, 227, 680-685), and protein concentrations were calculated from the absorbance at 280 nm,  $\epsilon_{280}^{0.1\%} = 1.17$  (Maturbara, H., et al. (1965), J. Biol. Chem., 240, 1125-1130).

35

Enzyme activity was measured with 200 $\mu$ g/mL succinyl-L-AlaL-AlaL-ProL-Phep-nitroanilide (Sigma) in 0.1M Tris pH 8.6 or 0.1 M CAPS pH 10.8 at 25°C. Specific activity ( $\mu$  moles product/min-mg) was calculated from the change in absorbance at 410 nm from production of p-nitroaniline with time per mg of enzyme (E410 = 8,480 M-lcm-l; Del Mar, E.G., et al. (1979), Anal. Biochem., 99, 316-320). Alkaline autolytic stability studies were performed on purified enzymes (200 $\mu$ g/mL) in 0.1 M potassium phosphate (pH 12.0) at 37°C. At various times aliquots were assayed for residual enzyme activity (Wells, J.A., et al. (1986) J. Biol. Chem., 261, 6564-6570).

## 15 E. Results

### 1. Optimization and analysis of mutagenesis frequency

A set of primer-template molecules that were randomly 3'-terminated over the subtilisin gene (Fig. 31) was produced by variable extension from a fixed 5'-primer (The primer mutated a unique AvaI site at codon 11 in the subtilisin gene). This was achieved by stopping polymerase reactions with EDTA after various times of extension. The extent and distribution of duplex formation over the 1 kb subtilisin gene fragment was assessed by multiple restriction digestion (not shown). For example, production of new HinfI fragments identified when polymerase extension had proceeded past Ile110, Leu233, and Asp259 in the subtilisin gene.

Misincorporation of each dNTPs at randomly terminated 3' ends by AMV reverse transcriptase (Zakour, R.A., et al. (1982), Nature, 295, 708-710; Zakour, R.A., et al. (1984), Nucleic Acids Res., 12, 6615-6628), used

conditions previously described (Champoux, J.J., (1984), Genetics, 2, 454-464). The efficiency of each misincorporation reaction was estimated to be greater than 80% by the addition of each dNTP's to the AvaI restriction primer, and analysis by polyacrylamide gel electrophoresis. Misincorporations were sealed by polymerization with all four dNTP's and closed circular DNA was produced by reaction with DNA ligase.

Several manipulations were employed to maximize the yield of the mutant sequences in the heteroduplex. These included the use of a deoxyuridine containing template (Kunkel, T.A. (1985), Proc. Natl. Acad. Sci. USA, 82 488-492; Pukkila, P.J. et al. (1983), Genetics, 104, 571-582), in vitro methylation of the mutagenic strand (Kramer, W. et al. (1982) Nucleic Acids Res., 10 6475-6485), and the use of AvaI restriction-selection against the wild-type template strand which contained a unique AvaI site. The separate contribution of each of these enrichment procedures to the final mutagenesis frequency was not determined, except that prior to AvaI restriction-selection roughly one-third of the segregated clones in each of the four pools still retained a wild-type AvaI site within the subtilisin gene. After AvaI restriction-selection greater than 98% of the plasmids lacked the wild-type AvaI site.

The 1.5 kb EcoRI-BamHI subtilisin gene fragment that was resistant to AvaI restriction digestion, from each of the four CsCl purified M13 RF pools was isolated on low melting agarose. The fragment was ligated in situ from the agarose with a similarly cut E. coli-B. subtilis shuttle vector, pB0180, and transformed directly into E. coli LE392. Such direct ligation and transformation of DNA isolated from agarose avoided

loses and allowed large numbers of recombinants to be obtained (>100,000 per  $\mu$ g equivalent of input M13 pool).

5 The frequency of mutagenesis for each of the four dNTPs misincorporation reactions was estimated from the frequency that unique restriction sites were eliminated (Table XX). The unique restriction sites chosen for this analysis, ClaI, PvuII, and KpnI, were distributed over the subtilisin gene starting at codons 35, 104, and 166, respectively. As a control, 10 the mutagenesis frequency was determined at the PstI site located in the  $\beta$  lactamase gene which was outside the window of mutagenesis. Because the absolute mutagenesis frequency was close to the percentage of undigested plasmid DNA, two rounds of restriction- 15 selection were necessary to reduce the background of surviving uncut wild-type plasmid DNA below the mutant plasmid (Table XX). The background of surviving plasmid from wild-type DNA probably represents the sum total of spontaneous mutations, uncut wild-type 20 plasmid, plus the efficiency with which linear DNA can transform E. coli. Subtracting the frequency for unmutagenized DNA (background) from the frequency for mutant DNA, and normalizing for the window of mutagenesis sampled by a given restriction analysis 25 (4-6 bp) provides an estimate of the mutagenesis efficiency over the entire coding sequence (~1000 bp).

30

35

TABLE XX

0251446

	$\alpha$ -thiol dNTP misincor- porated <sup>(b)</sup>	Restriction Site Selection	% resistant clones <sup>c</sup>			% resistant clones over Background <sup>d</sup>	% mutants per 1000bp <sup>e</sup>
			1st round	2nd round	Total		
5	None	<u>PstI</u>	0.32	0.7	0.002	0	-
	G	<u>PstI</u>	0.33	1.0	0.003	0.001	0.2
	T	<u>PstI</u>	0.32	<0.5	<0.002	0	0
	C	<u>PstI</u>	0.43	3.0	0.013	0.011	3
10	None	<u>ClaI</u>	0.28	5	0.014	0	-
	G	<u>ClaI</u>	2.26	85	1.92	1.91	380
	T	<u>ClaI</u>	0.48	31	0.15	0.14	35
	C	<u>ClaI</u>	0.55	15	0.08	0.066	17
15	None	<u>PvuII</u>	0.08	29	0.023	0	-
	G	<u>PvuII</u>	0.41	90	0.37	0.35	88
	T	<u>PvuII</u>	0.10	67	0.067	0.044	9
	C	<u>PvuII</u>	0.76	53	0.40	0.38	95
20	None	<u>KpnI</u>	0.41	3	0.012	0	-
	G	<u>KpnI</u>	0.98	35	0.34	0.33	83
	T	<u>KpnI</u>	0.36	15	0.054	0.042	8
	C	<u>KpnI</u>	1.47	26	0.38	0.37	93

25

(a) Mutagenesis frequency is estimated from the frequency for obtaining mutations that alter unique restriction sites within the mutagenized subtilisin gene (i.e., ClaI, PvuII, or KpnI) compared to mutation frequencies of the PstI site, that is outside the window of mutagenesis.

30

(b) Plasmid DNA was from wild-type (none) or mutagenized by dNTPs misincorporation as described.

(c) Percentage of resistant clones was calculated from the fraction of clones obtained after three fold or greater over-digestion of the plasmid with the indicated restriction enzyme compared to a

35

5 non-digested control. Restriction-resistant plasmid DNA from the first round was subjected to a second round of restriction-selection. The total represents the product of the fractions of resistant clones obtained from both rounds of selection and gives percentage of restriction-site mutant clones in the original starting pool. Frequencies were derived from counting at least 20 colonies and usually greater than 100.

(d) Percent resistant clones was calculated by subtracting the percentage of restriction-resistant clones obtained for wild-type DNA (i.e., none) from that obtained for mutant DNA.

10 (e) This extrapolates from the frequency of mutation over each restriction site to the entire subtilisin gene (~1 kb). This has been normalized to the number of possible bases (4-6 bp) within each restriction site that can be mutagenized by a given misincorporation event.

15

From this analysis, the average percentage of subtilisin genes containing mutations that result from dGTPas, dCTPAs, or dTTPAs misincorporation was estimated to be 90, 70, and 20 percent, respectively. These high mutagenesis frequencies were generally quite variable depending upon the dNTPas and misincorporation efficiencies at this site. Misincorporation efficiency has been reported to be both dependent on the kind of mismatch, and the context of primer (Champoux, J.J., (1984); Skinner, J.A., et al. (1986) Nucleic Acids Res., 14, 6945-6964). Biased misincorporation efficiency of dGTPas and dCTPAs over dTTPAs has been previously observed (Shortle, D., et al. (1985), Genetics, 110, 539-555). Unlike the dGTPas, dCTPAs, and dTTPAs libraries the efficiency of mutagenesis for the dATPAs

35

misincorporation library could not be accurately assessed because 90% of the restriction-resistant plasmids analyzed simply lacked the subtilisin gene insert. This problem probably arose from self-ligation of the vector when the dATP<sub>as</sub> mutagenized subtilisin gene was subcloned from M13 into pB0180. Correcting for the vector background, we estimate the mutagenesis frequency around 20 percent in the dATP<sub>as</sub> misincorporation library. In a separate experiment (not shown), the mutagenesis efficiencies for dGTP<sub>as</sub> and dTTP<sub>as</sub> misincorporation were estimated to be around 50 and 30 percent, respectively, based on the frequency of reversion of an inactivating mutation at codon 169.

The location and identity of each mutation was determined by a single track of DNA sequencing corresponding to the misincorporated  $\alpha$ thiodeoxy-nucleotide over the entire gene followed by a complete four track of DNA sequencing focused over the site of mutation. Of 14 mutants identified, the distribution was similar to that reported by Shortle and Lin (1985) except we did not observe nucleotide insertion or deletion mutations. The proportion of AG mutations was highest in the G misincorporation library, and some unexpected point mutations appeared in the dTTP<sub>as</sub> and dCTP<sub>as</sub> libraries.

## 2. Screening and Identification of Alkaline Stability Mutants of Subtilisin

It is possible to screen colonies producing subtilisin by halos of casein digestion (Wells, J.A. et al. (1983) Nucleic Acids Res., 11, 7911-7925). However, two problems were posed by screening colonies under high alkaline conditions (>pH 11). First, B. subtilis



will not grow at high pH, and we have been unable to transform an alkylphilic strain of bacillus. This problem was overcome by adopting a replica plating strategy in which colonies were grown on filters at neutral pH to produce subtilisin and filters subsequently transferred to casein plates at pH 11.5 to assay subtilisin activity. However, at pH 11.5 the casein micells no longer formed a turbid background and thus prevented a clear observation of proteolysis halos. The problem was overcome by briefly staining the plate with Coomassie blue to amplify proteolysis zones and acidifying the plates to develop casein micell turbidity. By comparison of the halo size produced on the reference growth plate (pH 7) to the high pH plate (pH 11.5), it was possible to identify mutant subtilisins that had increased (positives) or decreased (negatives) stability under alkaline conditions.

Roughly 1000 colonies were screened from each of the four misincorporation libraries. The percentage of colonies showing a differential loss of activity at pH 11.5 versus pH 7 represented 1.4, 1.8, 1.4, and 0.6% of the total colonies screened from the thiol dGTPas, dATPas, dTTPas, and dCTPAs libraries, respectively. Several of these negative clones were sequenced and all were found to contain a single base change as expected from the misincorporation library from which they came. Negative mutants included A36, E170 and V50. Two positive mutants were identified as V107 and R213. The ratio of negatives to positives was roughly 50:1.

### 3. Stability and Activity of Subtilisin Mutants at Alkaline pH

Subtilisin mutants were purified and their autolytic stabilities were measured by the time course of inactivation at pH 12.0 (Figs. 32 and 33). Positive mutants identified from the screen (i.e., V107 and R213) were more resistant to alkaline induced autolytic inactivation compared to wild-type; negative mutants (i.e., E170 and V50) were less resistant. We had advantageously produced another mutant at position 50 (F50) by site-directed mutagenesis. This mutant was more stable than wild-type enzyme to alkaline autolytic inactivation (Fig. 33) At the termination of the autolysis study, SDS-PAGE analysis confirmed that each subtilisin variant had autolyzed to an extent consistent with the remaining enzyme activity.

The stabilizing effects of V107, R213, and F50 are cumulative. See Table XXI. The double mutant, V107/R213 (made by subcloning the 920 bp EcoRI-KpnI fragment of pB0180V107 into the 6.6 kb EcoRI-KpnI fragment of pB0180R213), is more stable than either single mutant. The triple mutant, F50/V107/R213 (made by subcloning the 735 bp EcoRI-PvuII fragment of pF50 (Example 2) into the 6.8 kb EcoRI-PvuII fragment of pB0180/V107, is more stable than the double mutant V107/R213 or F50. The inactivation curves show a biphasic character that becomes more pronounced the more stable the mutant analyzed. This may result from some destabilizing chemical modification(s) (eg., deamidation) during the autolysis study and/or reduced stabilization caused by complete digestion of larger autolysis peptides. These alkaline autolysis studies have been repeated on separately purified enzyme batches with essentially the same results. Rates of autolysis should depend both on the conformational

F. Random Cassette Mutagenesis  
of Residues 197 through 228

Plasmid pΔ222 (Wells, et al. (1985) Gene 34, 315-323) was digested with PstI and BamHI and the 0.4 kb PstI/BamHI fragment (fragment 1, see Fig. 34) purified from a polyacrylamide gel by electroelution.

The 1.5 kb EcoRI/BamHI fragment from pS4.5 was cloned into M13mp9. Site directed mutagenesis was performed to create the A197 mutant and simultaneously insert a silent SstI site over codons 195-196. The mutant EcoRI/BamHI fragment was cloned back into pBS42. The pA197 plasmid was digested with BamHI and SstI and the 5.3 kb BamHI/SstI fragment (fragment 2) was purified from low melting agarose.

Complimentary oligonucleotides were synthesized to span the region from SstI (codons 195-196) to PstI (codons 228-230). These oligodeoxynucleotides were designed to (1) restore codon 197 to the wild type, (2) re-create a silent KpnI site present in pΔ222 at codons 219-220, (3) create a silent SmaI site over codons 210-211, and (4) eliminate the PstI site over codons 228-230 (see Fig. 35). Oligodeoxynucleotides were synthesized with 2% contaminating nucleotides at each cycle of synthesis, e.g., dATP reagent was spiked with 2% dCTP, 2% dGTP, and 2% dTTP. For 97-mers, this 2% poisoning should give the following percentages of non-mutant, single mutants and double or higher mutants per strand with two or more misincorporations per complimentary strand: 14% non-mutant, 28% single mutant, and 57% with ≥2 mutations, according to the general formula

$$f = \frac{\mu^n}{n!} e^{-\mu}$$

where  $\mu$  is the average number of mutations and  $n$  is a number class of mutations and  $f$  is the fraction of the total having that number of mutations. Complimentary oligodeoxynucleotide pools were phosphorylated and annealed (fragment 3) and then ligated at 2-fold molar excess over fragments 1 and 2 in a three-way ligation.

E. coli MM294 was transformed with the ligation reaction, the transformation pool grown up over night and the pooled plasmid DNA was isolated. This pool represented  $3.4 \times 10^4$  independent transformants. This plasmid pool was digested with PstI and then used to retransform E. coli. A second plasmid pool was prepared and used to transform B. subtilis (BG2036). Approximately 40% of the BG2036 transformants actively expressed subtilisin as judged by halo-clearing on casein plates. Several of the non-expressing transformants were sequenced and found to have insertions or deletions in the synthetic cassettes. Expressing BG2036 mutants were arrayed in microtiter dishes with 150  $\mu$ l of LB/12.5  $\mu$ g/mL chloramphenicol (cmp) per well, incubated at 37°C for 3-4 hours and then stamped in duplicate onto nitrocellulose filters laid on LB 1.5% skim milk/5  $\mu$ g/mL cmp plates and incubated overnight at 33°C (until halos were approximately 4-8 mm in diameter). Filters were then lifted to stacks of filter paper saturated with 1 x Tide commercial grade detergent, 50 mM Na<sub>2</sub>CO<sub>3</sub>, pH 11.5 and incubated at 65°C for 90 min. Overnight growth plates were Commassie stained and destained to establish basal levels of expression. After this treatment, filters were returned to pH7/skim milk/20  $\mu$ g/mL tetracycline plates and incubated at 37°C for 4 hours to overnight.

5 Mutants identified by the high pH stability screen to be more alkaline stable were purified and analyzed for autolytic stability at high pH or high temperature. The double mutant C204/R213 was more stable than wild type at either high pH or high temperature (Table XXII).

10 This mutant was dissected into single mutant parents (C204 and R213) by cutting at the unique SmaI restriction site (Fig. 35) and either ligating wild type sequence 3' to the SmaI site to create the single C204 mutant or ligating wild type sequence 5' to the SmaI site to create the single R213 mutant. Of the two single parents, C204 was nearly as alkaline stable as the parent double mutant (C04/R213) and slightly more thermally stable. See Table XXII. The R213 mutant was only slightly more stable than wild type under both conditions (not shown).

20 Another mutant identified from the screen of the 197 to 228 random cassette mutagenesis was R204. This mutant was more stable than wild type at both high pH and high temperature but less stable than C204.

25

30

35

TABLE XXII

Stability of subtilisin variants

5 Purified enzymes (200 $\mu$ g/mL) were incubated in 0.1M phosphate, pH 12 at 30°C for alkaline autolysis, or in  
 10 2mM CaCl<sub>2</sub>, 50mM MOPS, pH 7.0 at 62°C for thermal autolysis. At various times samples were assayed for residual enzyme activity. Inactivations were roughly pseudo-first order, and t 1/2 gives the time it took to reach 50% of the starting activity in two separate experiments.

15	Subtilisin variant	t 1/2 (alkaline autolysis)		t 1/2 (thermal autolysis)	
		Exp. #1	Exp. #2	Exp. #1	Exp. #2
	wild type	30	25	20	23
20	F50/V107/R213	49	41	18	23
	R204	35	32	24	27
	C204	43	46	38	40
	C204/R213	50	52	32	36
25	L204/R213	32	30	20	21

G. Random Mutagenesis at Codon 204

30 Based on the above results, codon 204 was targeted for random mutagenesis. Mutagenic DNA cassettes (for codon at 204) all contained a fixed R213 mutation which was found to slightly augment the stability of the C204 mutant.

35

Plasmid DNA encoding the subtilisin mutant C204/R213 was digested with SstI and EcoRI and a 1.0 kb EcoRI/SstI fragment was isolated by electro-elution from polyacrylamide gel (fragment 1, see Fig. 35).

5

C204/R213 was also digested with SmaI and EcoRI and the large 4.7 kb fragment, including vector sequences and the 3' portion of coding region, was isolated from low melting agarose (fragment 2, see Fig. 36).

10

Fragments 1 and 2 were combined in four separate three-way ligations with heterophosphorylated fragments 3 (see Figs. 36 and 37). This heterophosphorylation of synthetic duplexes should preferentially drive the phosphorylated strand into the plasmid ligation product. Four plasmid pools, corresponding to the four ligations, were restricted with SmaI in order to linearize any single cut C204/R213 present from fragment 2 isolation, thus reducing the background of C204/R213. E. coli was then re-transformed with SmaI-restricted plasmid pools to yield a second set of plasmid pools which are essentially free of C204/R213 and any non-segregated heterduplex material.

25

These second enriched plasmid pools were then used to transform B. subtilis (BG2036) and the resulting four mutant pools were screened for clones expressing subtilisin resistant to high pH/temperature inactivation. Mutants found positive by such a screen were further characterized and identified by sequencing.

30

The mutant L204/R213 was found to be slightly more stable than the wild type subtilisin. See Table XXII.

35

Having described the preferred embodiments of the present invention, it will appear to those ordinarily skilled in the art that various modifications may be made to the disclosed embodiments, and that such modifications are intended to be within the scope of  
5 the present invention.

10

15

20

25

30

35



CLAIMS;

1. A carbonyl hydrolase mutant having at least one property which is substantially different from the same property of a precursor carbonyl hydrolase from which the amino acid sequence of said carbonyl hydrolase mutant is derived, said property being selected from the group consisting of thermal stability and alkaline stability wherein said precursor carbonyl hydrolase is selected from the group consisting of naturally occurring carbonyl hydrolases and recombinant carbonyl hydrolases and said carbonyl hydrolase mutant amino acid sequence is derived by a method selected from the group consisting of the substitution, deletion and insertion of at least one amino acid in said amino acid sequence of said precursor carbonyl hydrolase.

2. A carbonyl hydrolase mutant having at least one property which is substantially different from the same property of a precursor carbonyl hydrolase from which the amino acid sequence of said carbonyl hydrolase mutant is derived, said property being selected from the group consisting of oxidative stability, substrate specificity, catalytic activity, thermal stability, alkaline stability and pH activity profile wherein said precursor carbonyl hydrolase is selected from the group consisting of naturally occurring carbonyl hydrolases and recombinant carbonyl hydrolases and said carbonyl hydrolase mutant amino acid sequence is derived by a method selected from the group consisting of deletion and insertion of at least one amino acid in said amino acid sequence of said precursor carbonyl hydrolase and substitution of more than one amino acid residue of said amino acid sequence of said precursor carbonyl hydrolase.

3. A carbonyl hydrolase mutant derived by the replacement of at least one amino acid residue of a precursor carbonyl hydrolase with a different amino acid, said one amino acid residue being selected from the group of amino acid residues of Bacillus amyloliquefaciens subtilisin consisting of Tyr21, Thr22, Ser24, Ser33, Asp36, Ala45, Gly46, Ala48, Ser49, Met50, Asn77, Ser87, Lys94, Val95, Leu96, Tyr104, Ile107, Gly110, Met124, Asn155, Glu156, Lys170, Tyr171, Pro172, Phe189, Asp197, Met199, Ser204, Lys213, Tyr217, Ser221, His67, Leu126, Leu135, Gly97, Asp99, Ser101, Gly102, Glu103, Leu126, Gly127, Gly128, Pro129, Tyr214, Gly215, and equivalent amino acid residues in other precursor carbonyl hydrolases.

4. A carbonyl hydrolase mutant having an amino acid sequence derived from the amino acid sequence of a precursor carbonyl hydrolase by the substitution of a different amino acid for more than one amino acid residue of said amino acid sequence of said precursor carbonyl hydrolase, said amino acid residues being selected from the group of amino acid residues of Bacillus amyloliquefaciens subtilisin consisting of Tyr21, Thr22, Ser24, Asp32, Ser33, Asp36, Ala45, Gly46, Ala48, Ser49, Met50, Asn77, Ser87, Lys94, Val95, Leu96, Tyr104, Ile107, Gly110, Met124, Ala152, Asn-155, Glu156, Gly166, Gly169, Lys170, Tyr171, Pro172, Phe189, Asp197, Met199, Ser204, Lys213, Tyr217, Ser221, Met222, His67, Leu126, Leu135, Gly97, Asp99, Ser101, Gly102, Glu103, Leu126, Gly127, Gly128, Pro129, Tyr214, Gly215, and equivalent amino acid residues in other precursor carbonyl hydrolases.

5. The mutant of Claim 4 wherein said combinations are selected from the group consisting of Thr22/Ser87, Ser24/Ser87, Ala45/Ala48, Ser49/Lys94, Ser49/Val95, Met50/Val95, Met50/Gly110, Met50/Met124, Met50/Met222, Met124/Met222, Glu156/Gly166, Glu156/Gly169, Gly166/Met222, Gly169/Met222, Tyr21/Thr22, Met50/Met124/Met222, Tyr21/Thr22/Ser87, Met50/Glu156/Gly166/Tyr217, Met50/Glu156/Tyr217, Glu156/Gly169/Tyr217, Ile170/Lys213, Ser204/Lys213, Met50/Ile107/Lys213 and Ser24/Met50/Ile107/Glu156/Gly166/Gly169/Ser204/Lys213/Gly215/Tyr217.

6. A carbonyl hydrolase mutant derived by the replacement of at least one amino acid residue of a precursor carbonyl hydrolase with a different amino acid, said one amino acid residue being selected from the group of amino acid residues of of Bacillus amyloliquefaciens subtilisin consisting of Tyr21, Thr22, Ser24, Asp32, Ser33, Asp36, Ala45, Gly46, Ala48, Ser49, Met50, Asn77, Ser87, Lys94, Val95, Leu96, Tyr104, Ile107, Gly110, Met124, Ala152, Asn-155, Glu156, Gly166, Gly169, Lys170, Tyr171, Pro172, Phe189, Asp197, Met199, Ser204, Lys213, Tyr217, Ser221, Met222, His67, Leu126, Leu135, Gly97, Asp99, Ser101, Gly102, Glu103, Leu126, Gly127, Gly128, Pro129, Tyr214, Gly215, and equivalent amino acid residues in other precursor carbonyl hydrolases, wherein said at least one amino acid residue of said precursor carbonyl hydrolase is replaced with the amino acid residues listed in TABLE I and TABLE II herein.

7. The mutant of Claim 6 wherein the amino acid replacing said at least one amino acid residue in said precursor carbonyl hydrolase is selected from the replacement amino acids listed in TABLE I herein.

8. Mutant DNA sequence encoding the mutant of claims  
1 through 7.

9. Expression vector containing the mutant DNA  
sequence of claim 8.

5

10. Host cell transformed with the expression vector  
of Claim 9.

10

15

20

25

30

35

0251446

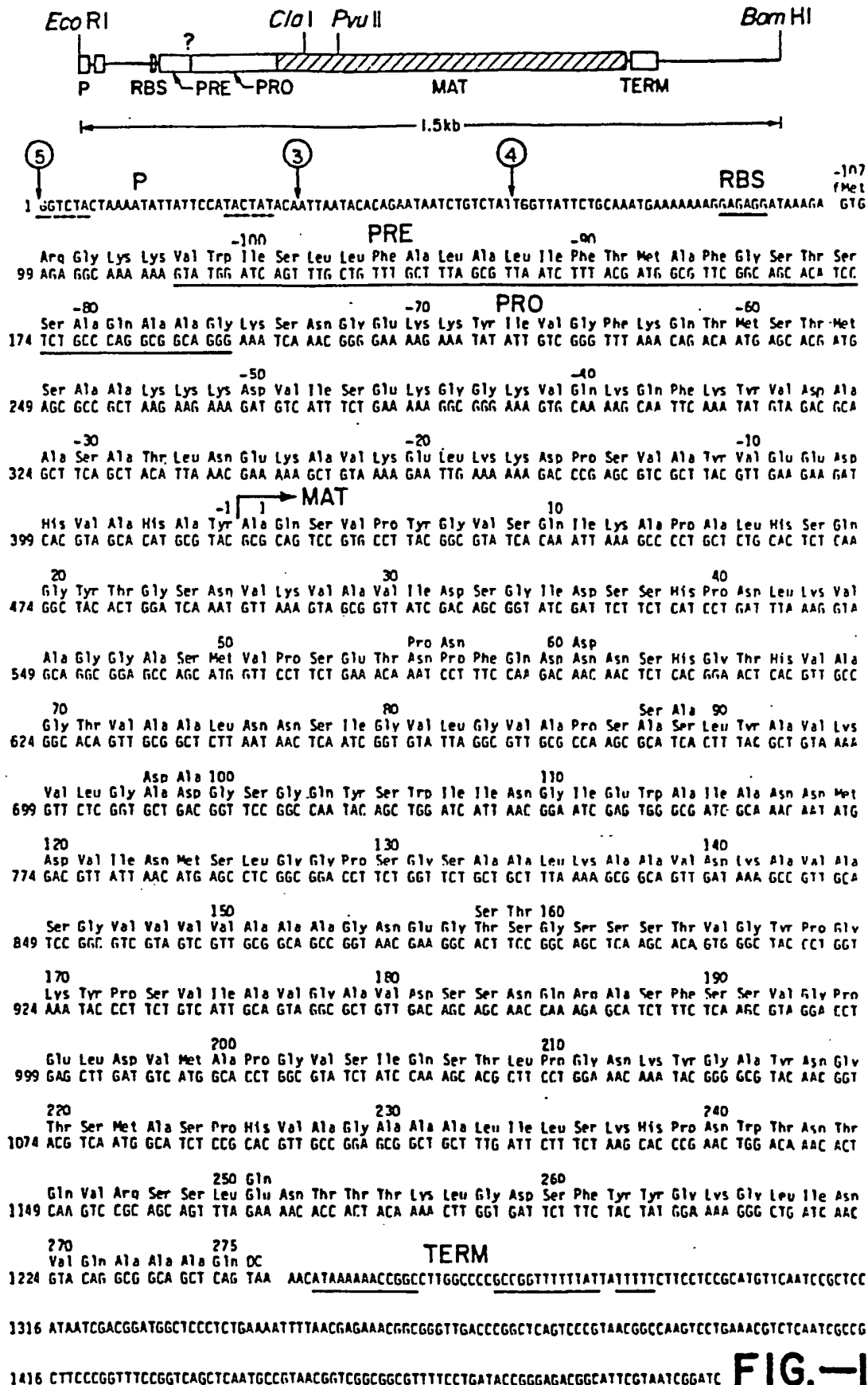
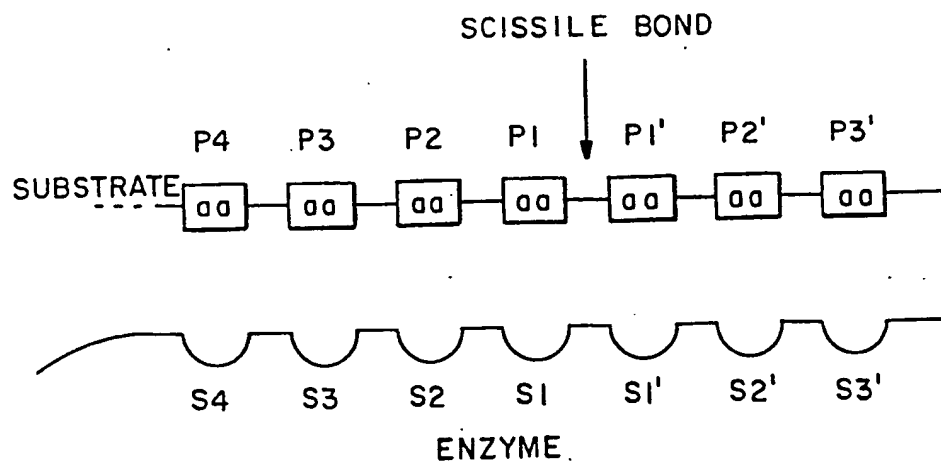
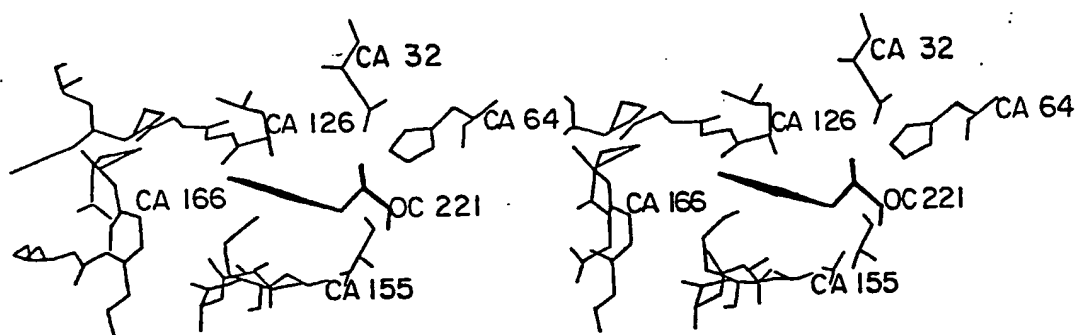


FIG.-1

**FIG. -2****FIG. - 3**

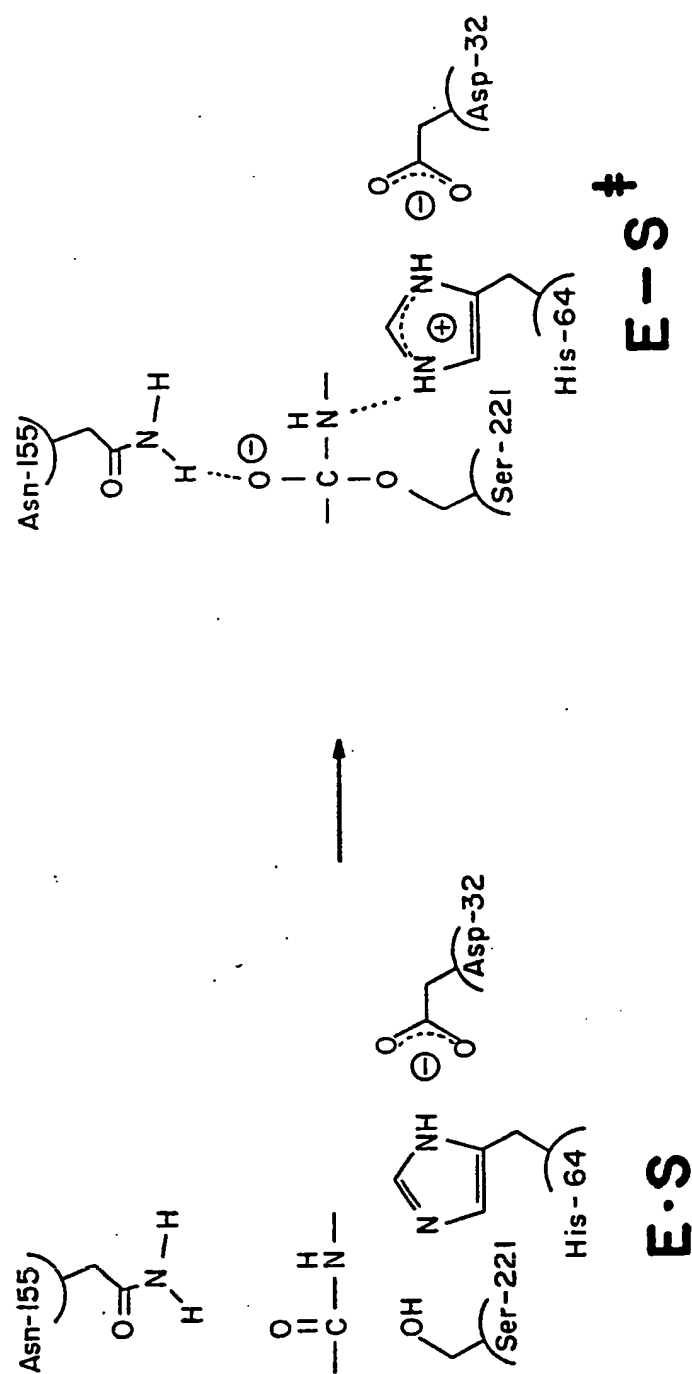


FIG.-4

Homology of *Bacillus proteases*

1. *Bacillus amyloliquifaciens*  
 2. *Bacillus subtilis* var. 1168  
 3. *Bacillus licheniformis* (carlsbergensis)

1	A	Q	S	V	P	Y	6	V	S	Q	I	K	A	P	A	L	H	S	Q	20
A	Q	S	V	P	Y	6	6	I	S	Q	I	K	A	P	A	L	H	S	Q	6
A	Q	T	V	P	Y	6	6	I	P	L	I	K	A	D	K	V	Q	A	Q	6
21										30										40
Y	T	6	S	N	V	K	V	A	V	V	I	D	S	6	I	D	S	S	H	P
Y	T	6	S	N	V	K	V	A	V	V	I	D	S	6	I	D	S	S	H	P
F	K	6	A	N	V	K	V	A	V	L	D	T	6	I	Q	A	S	H	P	
41										50										60
D	L	K	V	A	6	6	A	S	M	V	P	S	E	T	N	P	F	Q	D	
D	L	N	V	R	6	6	A	S	F	V	P	S	E	T	N	P	Y	Q	D	
D	L	N	V	V	6	6	A	S	F	V	A	6	E	A	Y	N	T	.	D	
61										70										80
N	N	S	H	6	T	H	V	A	6	T	V	A	A	L	N	N	S	I	6	
6	S	S	H	6	T	H	V	A	6	T	I	A	A	L	N	N	S	I	6	
6	N	6	H	6	T	H	V	A	6	T	V	A	A	L	D	N	T	T	6	
81										90										100
V	L	6	V	A	P	S	A	S	L	Y	A	V	K	V	L	6	A	D	6	
V	L	6	V	S	P	S	A	S	L	Y	A	V	K	V	L	D	S	T	6	
V	L	6	V	A	P	S	V	S	L	Y	A	V	K	V	L	N	S	S	6	
101										110										120
S	6	Q	Y	S	W	I	I	N	6	I	E	W	A	I	A	N	N	M	D	
S	6	Q	Y	S	W	I	I	N	6	I	E	W	A	I	S	N	N	M	D	
S	6	S	Y	S	6	I	V	S	6	I	E	W	A	T	T	N	6	M	D	

FIG.—5A-1



121	V	I	N	M	S	L	6	6	P	130	S	6	S	A	A	L	K	A	A	V	140	D
	V	I	N	M	S	L	6	6	P		T	6	S	T	A	L	K	T	V	V		D
	V	I	N	M	S	L	6	6	A		S	6	S	T	A	M	K	Q	A	V		D
141	K	A	V	A	S	6	V	V	V	150	V	A	A	A	6	N	E	6	T	S	160	6
	K	A	V	S	S	6	I	V	V		A	A	A	A	6	N	E	6	S	S		6
	N	A	Y	A	R	6	V	V	V		V	A	A	A	6	N	S	6	N	S		6
161	S	S	T	V	6	Y	P	6	170	K	Y	P	S	V	I	A	V	6	A	180	V	
	S	T	S	T	V	6	Y	P		K	Y	P	S	T	I	A	V	6	A		V	
	S	T	N	T	I	6	Y	P		K	Y	D	S	V	I	A	V	6	A		V	
181	D	S	S	N	Q	R	A	S	F	190	S	S	V	6	P	E	L	D	V	M	200	A
	N	S	S	N	Q	R	A	S	F		S	S	A	6	S	E	L	D	V	M		A
	D	S	N	S	N	R	A	S	F		S	S	V	6	A	E	L	E	V	M		A
201	P	6	V	S	I	Q	S	T	L	210	P	6	N	K	Y	6	A	Y	N	6	220	T
	P	6	V	S	I	Q	S	T	L		P	6	6	T	Y	6	A	Y	N	6		T
	P	6	A	6	V	Y	S	T	Y		P	T	N	T	Y	A	T	L	N	6		T
221	S	M	A	S	P	H	V	A	6	230	A	A	A	L	I	L	S	K	H	P	240	N
	S	M	A	T	P	H	V	A	6		A	A	A	L	I	L	S	K	H	P		T
	S	M	A	S	P	H	V	A	6		A	A	A	L	I	L	S	K	H	P		N
241	W	T	N	T	Q	V	R	S	S	250	L	E	N	T	T	T	K	L	6	D	260	S
	W	T	N	A	Q	V	R	D	R		L	E	S	T	A	T	Y	L	6	N		S
	L	S	A	S	Q	V	R	N	R		L	S	S	T	A	T	Y	L	6	S		S
261	F	Y	Y	6	K	6	L	I	N	270	V	Q	A	A	A	Q						
	F	Y	Y	6	K	6	L	I	N		V	Q	A	A	A	Q						
	F	Y	Y	6	K	6	L	I	N		V	E	A	A	A	Q						

FIG.—5A—2



A	A	A	G	N	E	S	T	S	150	G	S	S	S	T	V	G	Y	P	S	170	K
A	A	A	G	N	A	G	N	T	A	.	.	.	.	.	P	N	Y	P	A	Y	
Y	P	S	V	I	A	V	G	A	180	V	D	B	S	N	Q	R	A	S	F	180	S
Y	S	N	A	I	A	V	A	S	T	D	Q	N	D	N	K	S	S	F	S	S	
S	V	G	P	E	L	D	V	M	200	A	P	G	U	S	I	O	S	T	L	210	P
T	Y	G	S	V	V	D	V	A	A	P	G	S	U	I	Y	S	T	Y	P	P	
G	N	K	Y	G	A	J	N	G	220	T	S	M	A	S	P	H	V	A	G	230	A
T	S	T	Y	A	S	L	S	G	T	S	M	A	T	P	H	V	A	G	U	U	
A	A	L	I	L	S	K	H	P	240	N	U	T	N	T	Q	U	R	S	S	250	L
A	G	L	L	A	S	D	B	R	S	.	.	A	S	N	I	R	A	A	I	I	
E	N	T	T	T	K	.	L	G	260	D	S	F	Y	Y	G	K	G	L	I	N	
E	N	T	A	D	K	I	S	G	T	G	T	Y	U	A	K	B	R	V	N		
270	U	Q	A	A	A	D															
A	Y	K	A	U	Q	Y															

FIG.—5B-2

### TOTALLY CONSERVED RESIDUES IN SUBTILISINS

**FIG.—5C**

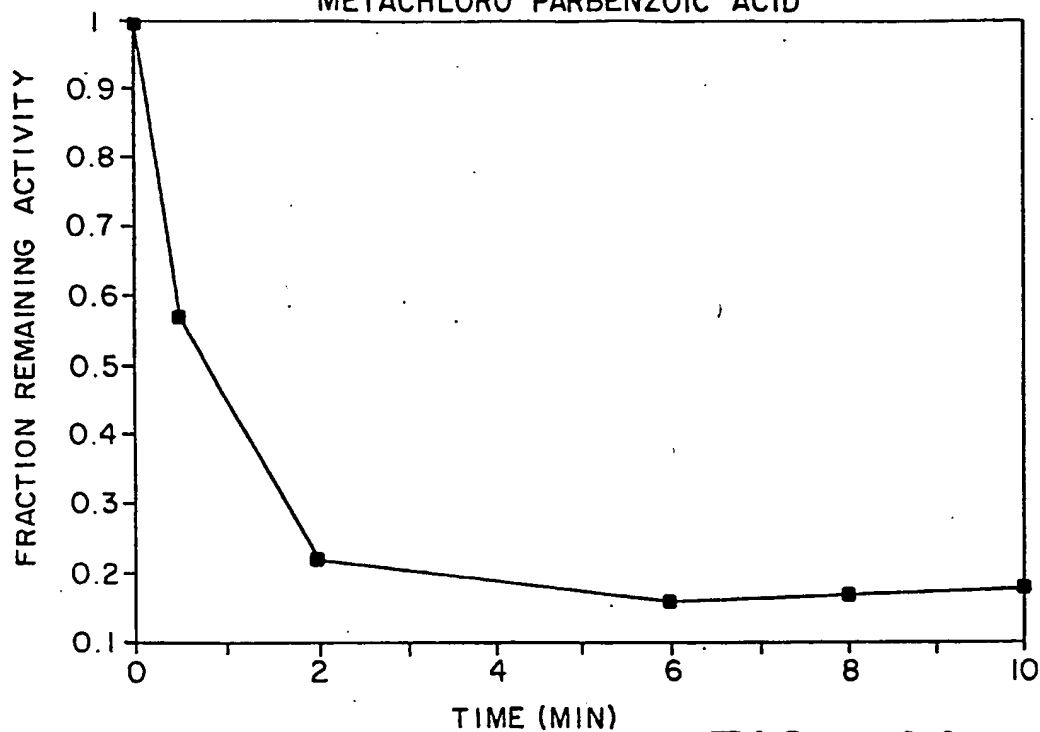
INACTIVATION OF L222 WITH  
METACHLORO PARBENZOIC ACID

FIG.-6A

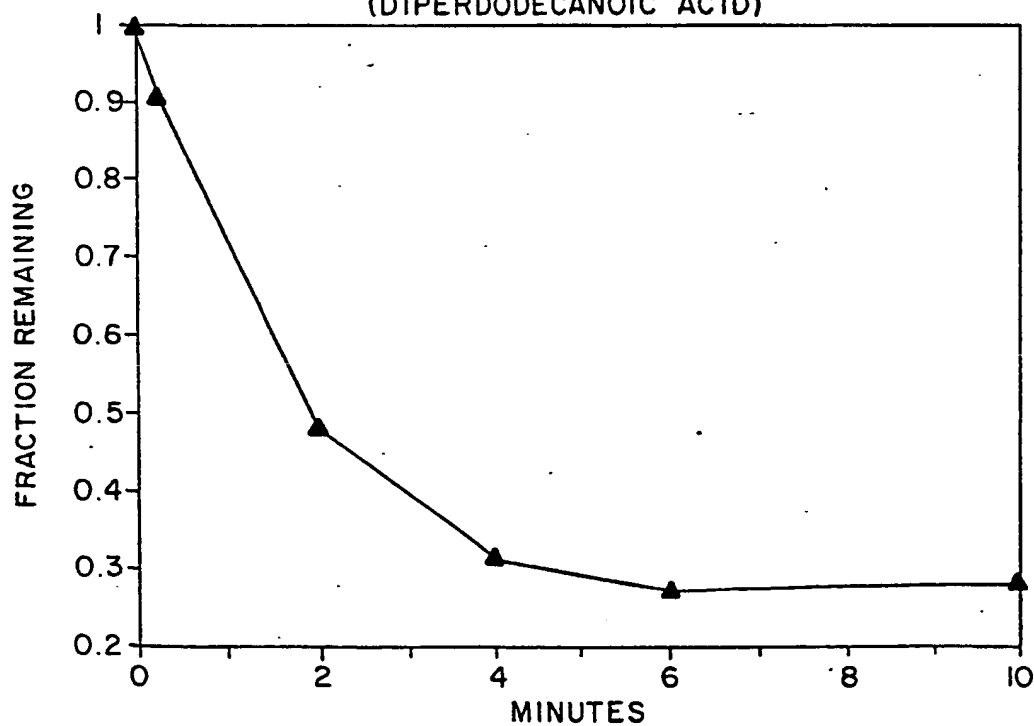
INACTIVATION OF Q222 BY DPDA  
(DIPERDODECANOIC ACID)

FIG.-6B

0251446

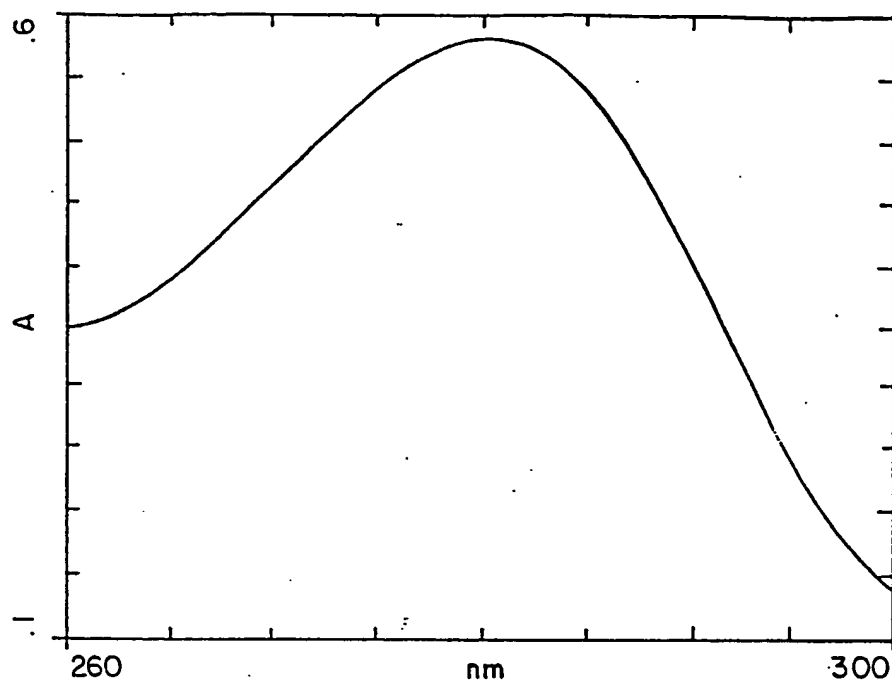


FIG. - 7A

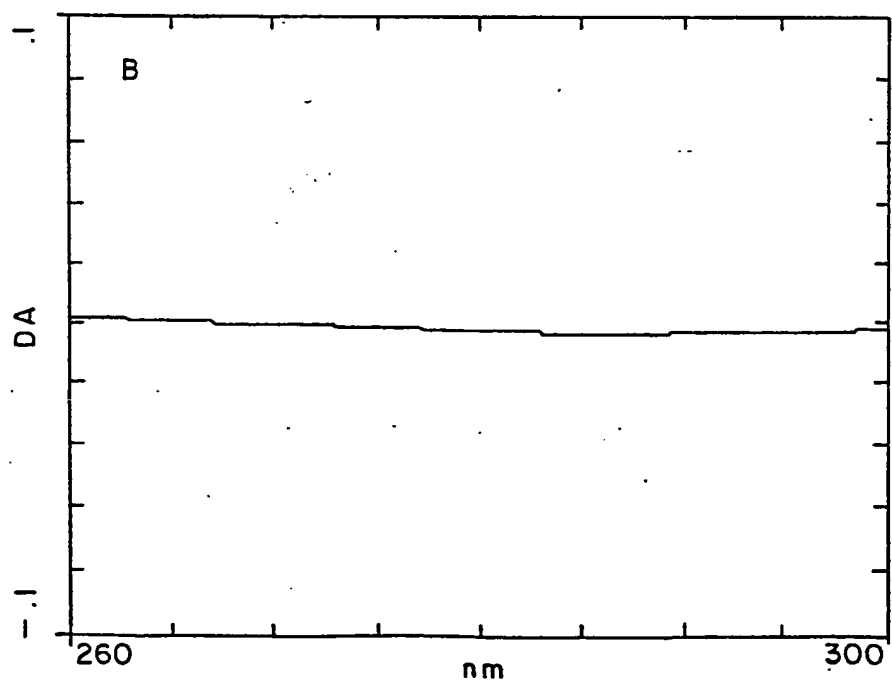


FIG. - 7B

0251446

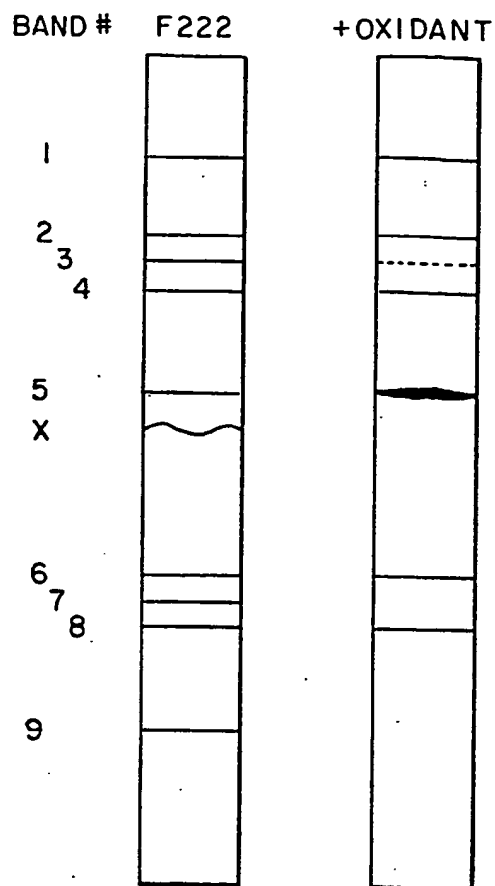


FIG.- 8

CNBr FRAGMENT MAP OF F222 MUTANT

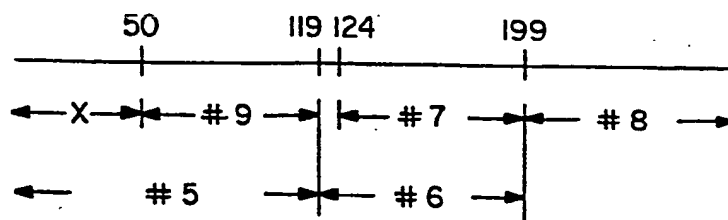


FIG.-9

1. Codon number: 43 45
2. Wild type amino acid sequence: Lys-Val-Ala-Gly-Gly-Ala-Ser-Met-Val-Pro-Ser
3. Wild type DNA sequence: 5'-AAG-GTA-GCA-GGC-GGA-GCC-AGC-ATG-GTT-CCT-TCT  
TTC-CAT-CGT-CCG-CCT-CCG-TCG-TAC-CAA-GGA-AGA-5'
4. pΔ50: 

\*\*\*

5'-AAG-GCC-T-----GC-ATG-GTA-CCT-TCT

TTC-CCG-A-----CG-TAC-CAT-GGA-AGA-5'

*Su* I

\*

5'-AAG-G-----GC-ATG-GTA-CCT-TCT

TTC-Cp
5. pΔ50 cut with *Su*I/*Kpn* I 

\*

5'-AAG-G-----GC-ATG-GTA-CCT-TCT

TTC-Cp

\*

5'-AAG-GTA-GCA-GGC-GGA-GCC-AGC-ATG-GTA-CCT-TCT

TTC-Cp
6. Cut pΔ50 ligated with cassettes: 

5'-AAG-GTA-GCA-GGC-GGA-GCC-AGC-ATG-GTA-CCT-TCT

TTC-Cp

5'-AAG-GTA-GCA-GGC-GGA-GCC-AGC-ATG-GTA-CCT-TCT

TTC-Cp
7. Mutagenesis primer for pΔ50: 

\*\*\*

5'-CT-GAT-TTA-AAG-GCC-TGC-ATG-GTA-CCT-TCT-GA

\*

5'-CT-GAT-TTA-AAG-GCC-TGC-ATG-GTA-CCT-TCT-GA
8. Mutants made: V45, P45, V45/P48, E46, E48, V48, C49, C50, F50

FIG.—10



1. Codon number: 117 120 124 126 130
2. Wild type amino acid sequence: Asn-Asn-Met-Asp-Val-Ile-Asn-Met-Ser-Leu-Gly-Gly-Pro-Ser
3. Wild type DNA sequence: 5'-AAC-AAT-ATG-GAC-GTT-ATT-AAC-ATG-AGC-CTC-GGC-GGA-CCT-TCT  
TTG-TTA-TAC-CTG-CAA-TAA-TTG-TAC-TCG-GAG-CCG-CCT-GGA-AGA-5'
4. pΔ124:
 

\*\*\*

5'-AAC-AAT-ATG-GAT-ATC-----C-GGC-GGC-CCT-TCT

TTG-TTA-TAC-CTA-TAG-----G-CCC-CCG-GGA-AGA-5'

Eco RV                      *Apa* I
5. pΔ124 cut with *Eco* RV and *Apa* I
 

\*

5'-AAC-AAT-ATG-GAT

TTG-TTA-TAC-CTAP
6. Cut pΔ124 ligated with cassettes:
 

\*

5'-AAC-AAT-ATG-GAT-GTT-ATT-AAC-ATG-AGC-CTC-GGC-GGC-CCT-TCT

TTG-TTA-TAC-CTA-CAA-TAA-TTG-TAC-TCG-GAG-CCG-CCT-GGA-AGA-5'
7. Mutagenesis primer for pΔ124::
 

\*\*\*

5'-AAC-AAT-ATG-GAT-ATC-C-GGC-GGC-CCT-TCT-GGT-TC-3'
8. Mutants made: I 124, L 124 AND C 126

FIG.—II

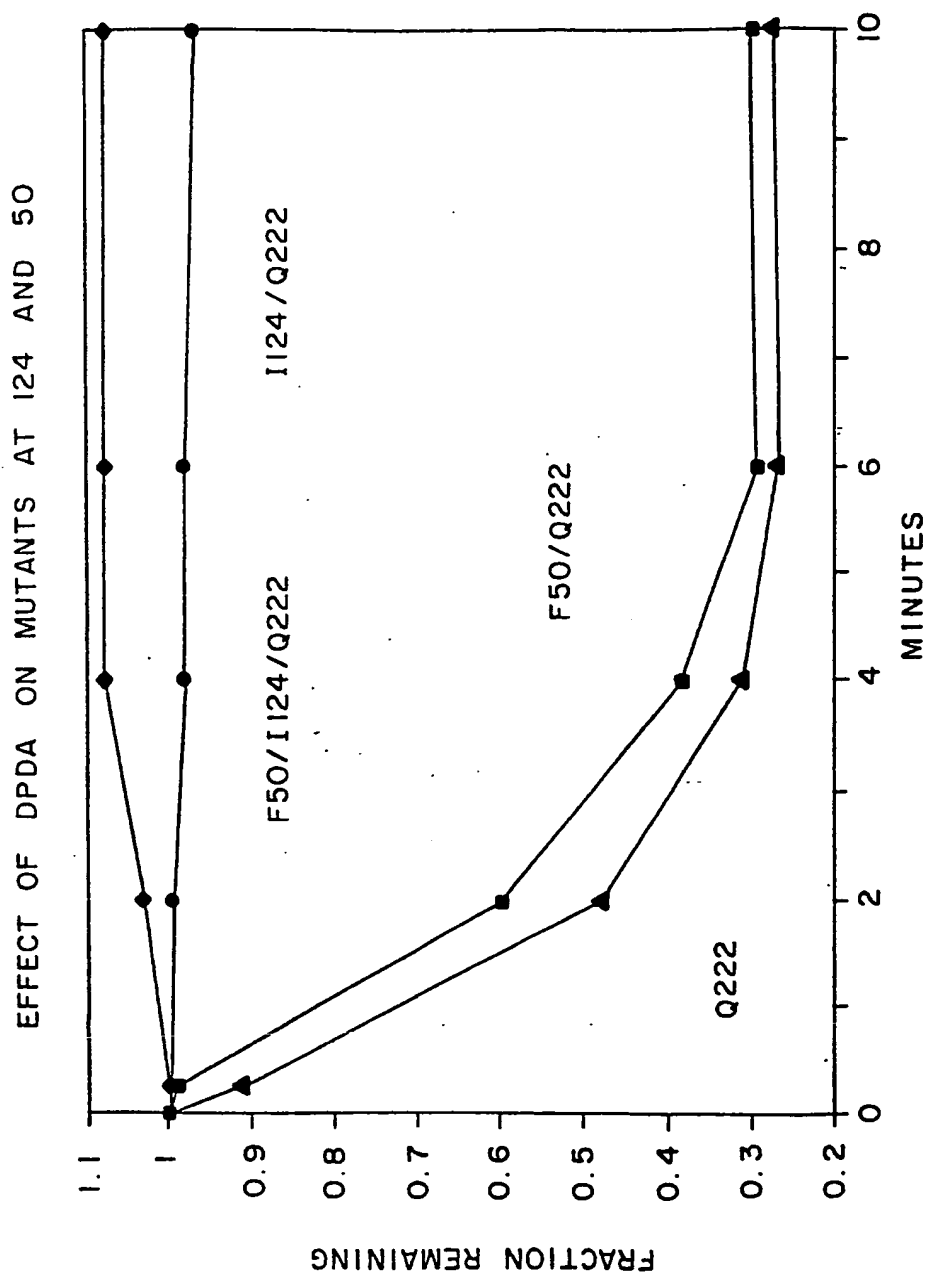


FIG.-12

**MUTAGENESIS PRIMER 37 MER**

5' AA GGC ACT TCC GGG AGC TCA ACC CGG GTA AA TAC CCT 3'

**FIG. -13**

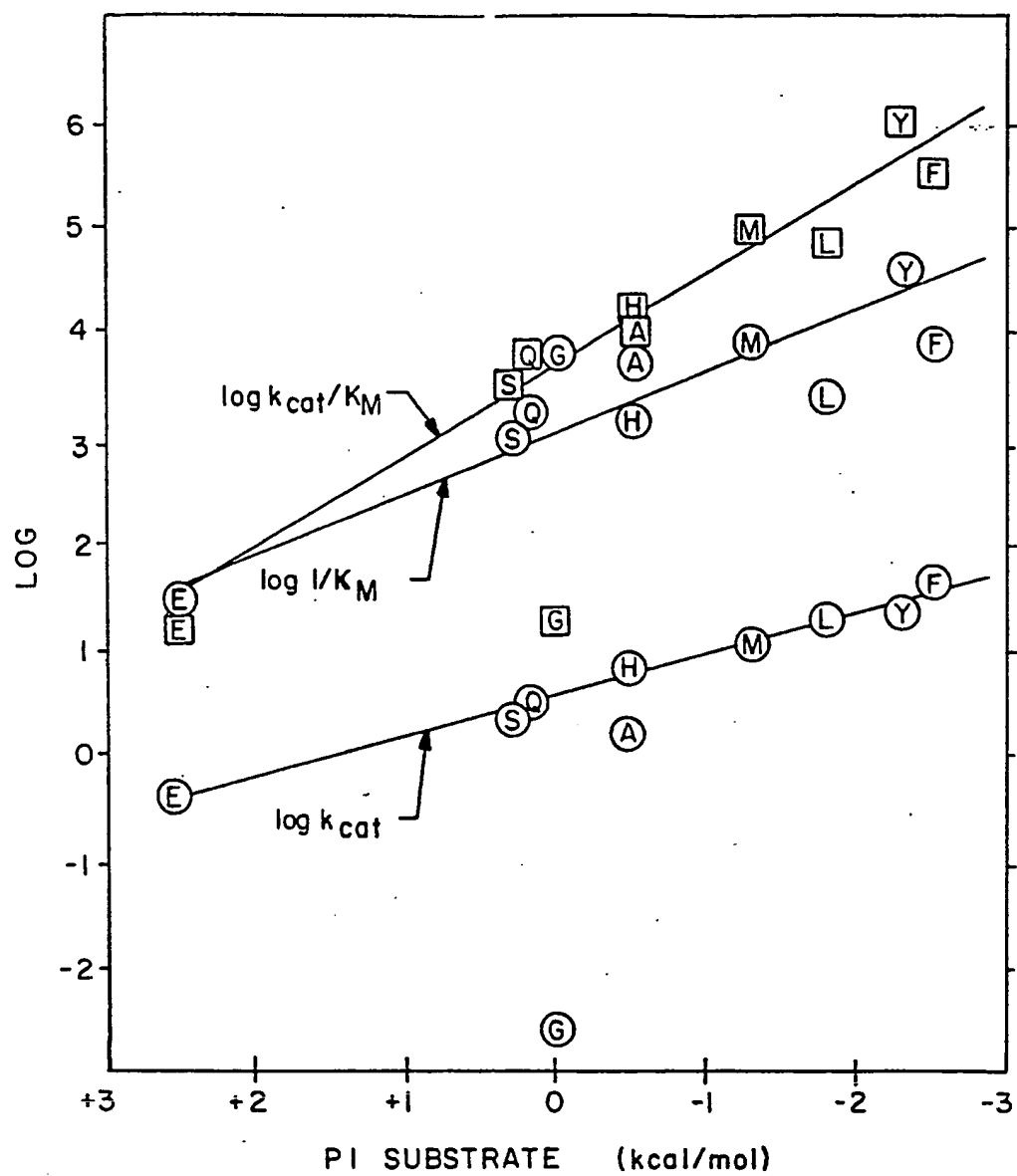


FIG. - 14

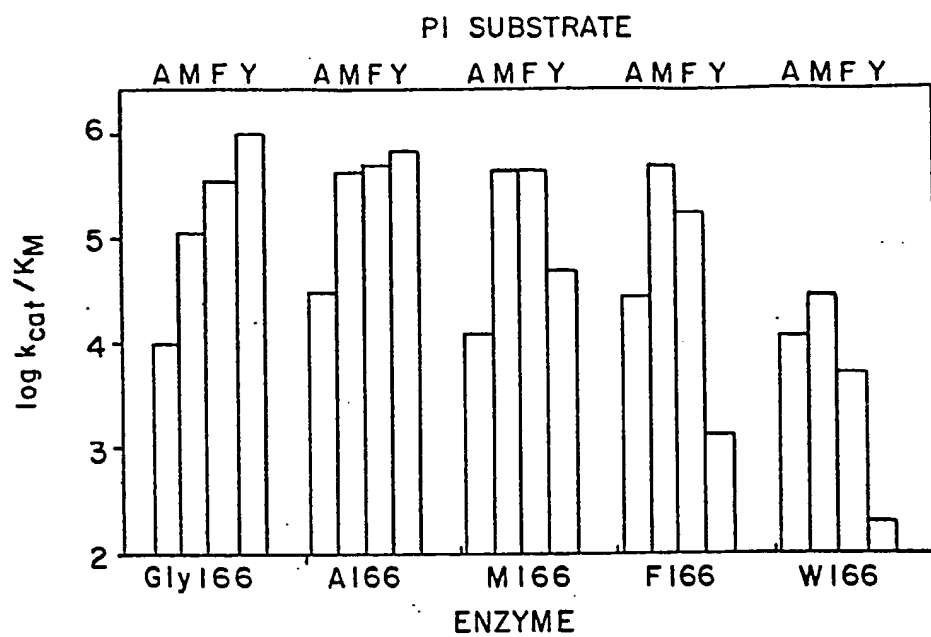


FIG.-15A

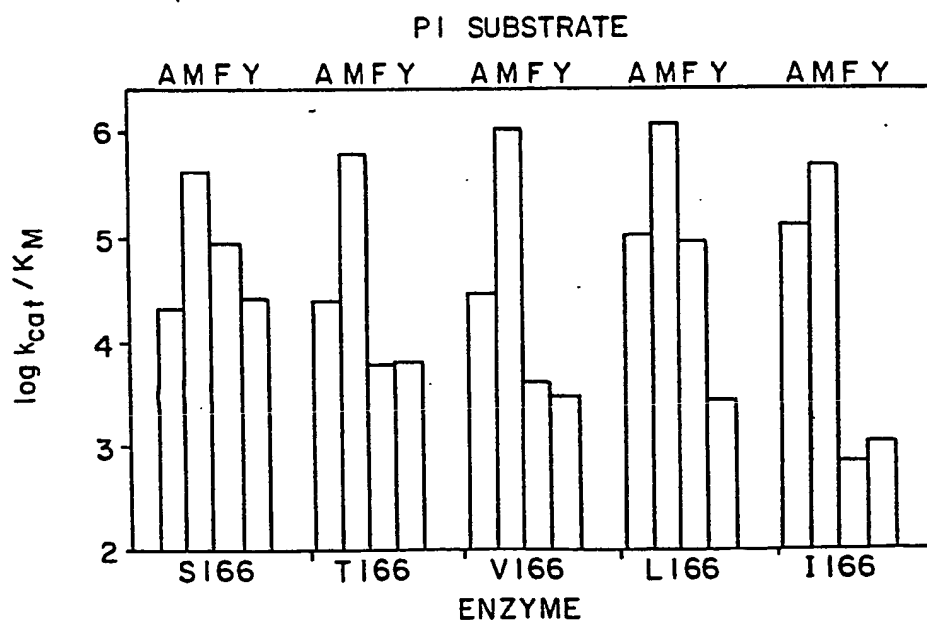


FIG.-15B

0251446

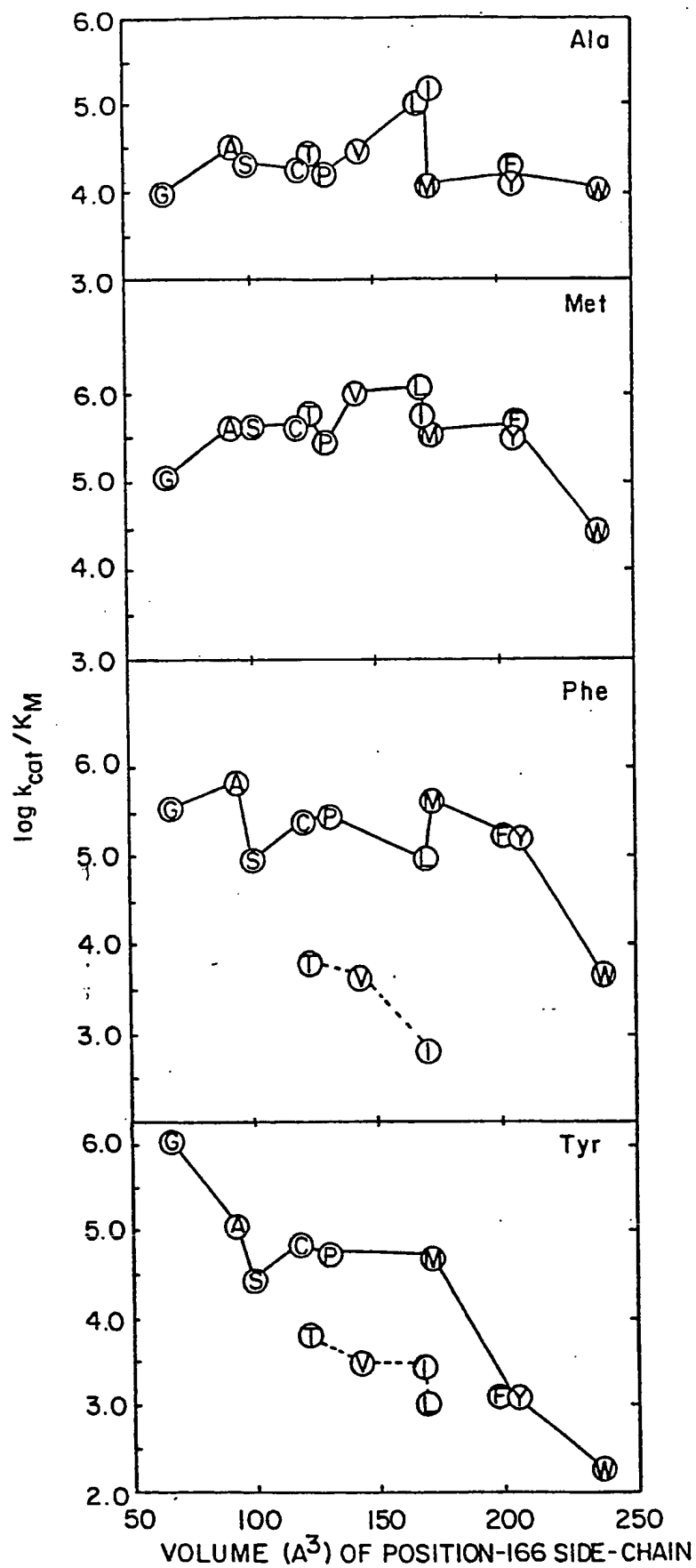


FIG.-16

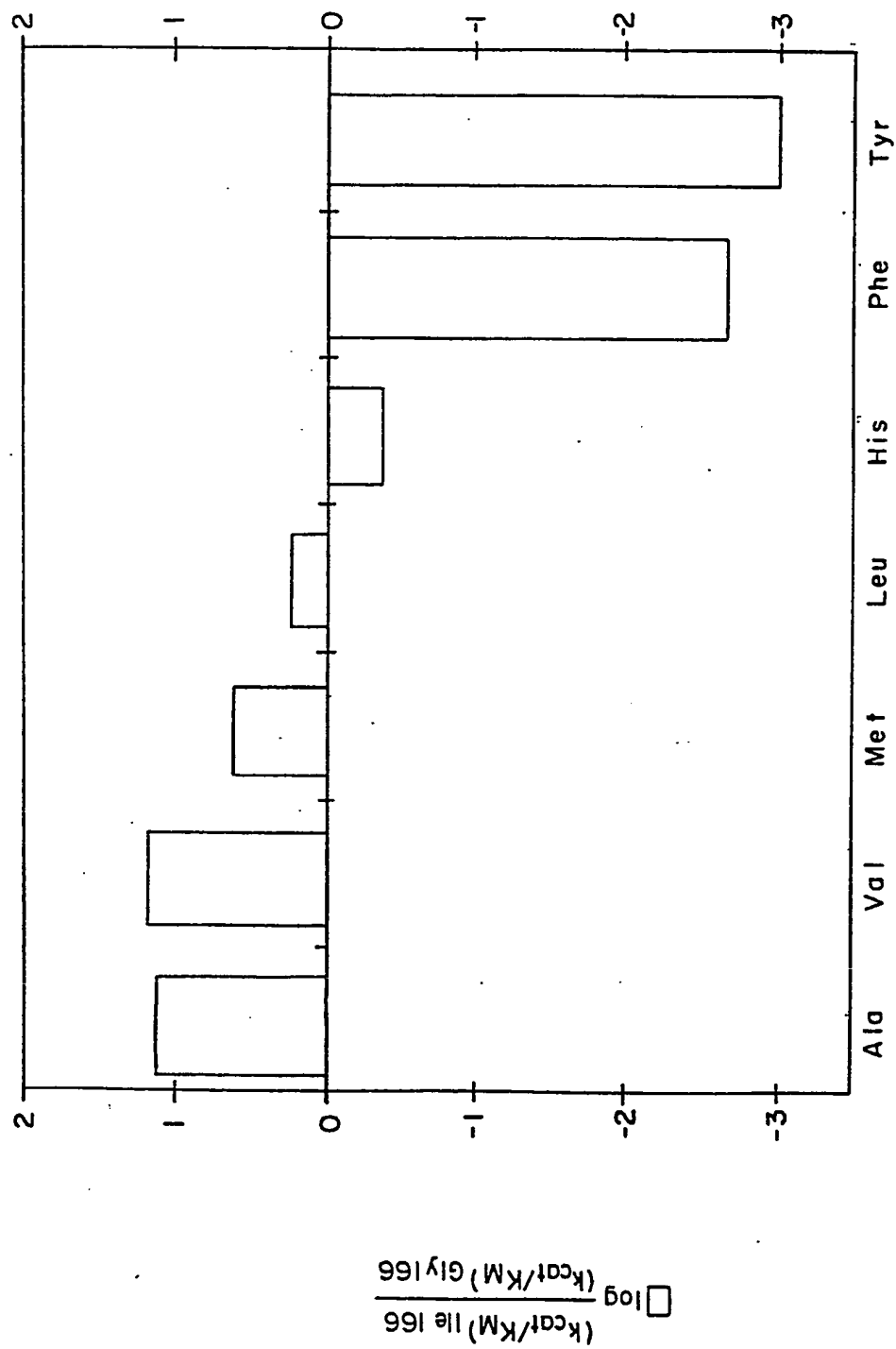


FIG. - 17

## GLY-169 CASSETTE MUTAGENESIS

WILD TYPE AMINO ACID SEQUENCE:	CODON:	
	162	169
	SER SER THR VAL GLY TYR PRO GLY LIS TYR PRO SER	173
1. WILD TYPE DNA SEQUENCE	5' TCA AGC ACA GTG GGC TAC CCT GGT AAA TAC CCT TCT 3'	
	3' AGT TCG TGT CAC CCG ATG GGA CCA TTT ATG GGA AGA 5'	
2. P169 DNA SEQUENCE	5' TCA AGC ACA GTC GGG TAC CCT-----GA TAT CCT TCT 3'	
	3' AGT TCG TGT CAC GCC ATG GGA CT ATA GGA AGA 5'	
	KPN I E C O R V	
3. P169 CUT WITH KPN I AND E C O R V:	5' TAC AGC ACA GTC GGG TAC PAT CCT TCT 3'	
	3' AGT TCG TGT CAC CCP TA GGA AGA 5'	
4. CUT P169 LIGATED WITH OLIGONUCLEOTIDE POOLS	5' TAC AGC ACA GTG GGG TAC CCT NNN AAA TAT CCT TGT 3'	
	3' AGT TCG TGT CAC CCC ATG GGA NNN TTT ATA GGA AGA 5'	
MUTAGENESIS PRIMER FOR P169	5' AAG CAC AGT GGG GTA CCC TGA TAT CCT TCT GTC A 3'	

FIG.—18



1. Codon number: 100 104 105 108
2. Wild type amino acid sequence: Gly-Ser-Gly-Gln-Tyr-Ser-Trp-Ile-Ile-
3. Wild type DNA sequence: 5'-GGT-TCC-GGC-CAA-TAC-AGC-TGG-ATC-ATT-3'  
Pvu I
4. Primer for *Hind* III  
Insertion at 104: 5'-GGT-TCC-GGC-CAA-GCTT-AGC-TGG-ATC-ATT-3'  
Hind III
5. Primers for 104 mutants: 5'-----T-TCC-GCC-CAA-NNN-AGC-TGG-ATC-----3'  
\*\*\*
6. Mutants made: A, M, L, S, AND H104

FIG.—19

1. Codon number: 148 150 152 155
2. Wild type amino acid sequence: Val-Val-Val-Ala-Ala-Ala-Gly-Asn-Glu
3. Wild type DNA sequence: 5'-GTA-GTC-GTT-GCG-GCA-GCC-GGT-AAC-GAA-3'
4. V152/P153 5'-GTA-GTC-GTT-GCG-GTA-CCG-GGT-AAC-GAA-3'  

$\begin{array}{c} * \quad * \\ \boxed{\text{GTA-CCG}} \\ \text{Kpn I} \end{array}$
5. S152: 5'-GTA-GTC-GTT-GCG-AGC-GCC-GGT-AAC-GAA-3'  

$\begin{array}{c} *** \end{array}$
6. G152: 5'-GTA-GTC-GTT-GCG-GGC-GCC-GGT-AAC-GAA-3'  

$\begin{array}{c} ** \end{array}$

FIG.—20

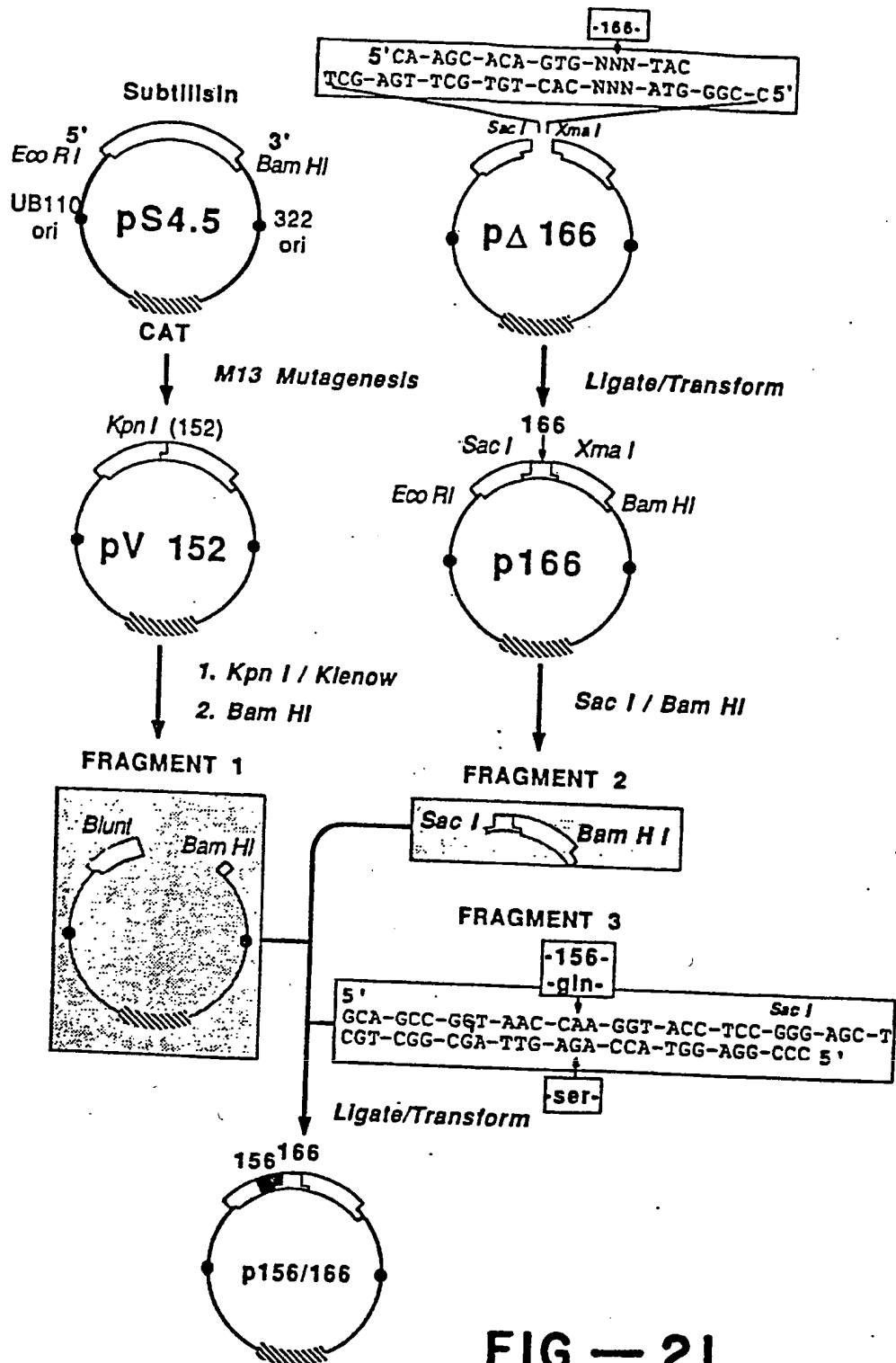


FIG.— 21

1. Codon number: 211 215 217 220
2. Wild type amino acid sequence: Gly-Asn-Lys-Tyr-Gly-Ala-Tyr-Asn-Gly-Thr-Ser-Met-Ala
3. Wild type DNA sequence: 5'-GGA-AAC-AAA-TAC-GGG-GCG-TAC-AAC-GGT-ACG-TCA-ATG-GCA  
CCT-TTG-TTT-ATG-CCC-CGC-ATG-TTG-CCA-TGC-AGT-TAC-CGT-5'
4. pΔ217  
5'-GGA-AAC-AAA-TAC-GGC-GCC-TAC-----GG-ATA-TCA-ATG-GCA  
CCT-TTG-TTT-ATG-CCG-CGG-ATG-----CC-TAT-AGT-TAC-CGT-5'  
Nar I Eco RV
5. pΔ217 cut with Nar I and Eco RI  
5'-GGA-AAC-AAA-TAC-GG\*  
CCT-TTG-TTT-ATG-CCG-Gp  
\* PA-TCA-ATG-GCA  
T-AGT-TAC-CGT-5'
6. Cut pΔ217 ligated with cassettes:  
5'-GGA-AAC-AAA-TAC-GGC-GCG-NNN-AAC-GGT-ACA-TCA-ATG-GCA  
CCT-TTG-TTT-ATG-CCG-CGC-NNN-TTG-CCA-TGT-AGT-TAC-CGT-5'  
\*\*\*
7. Mutagenesis primer for pΔ217:  
5'-GA-AAC-AAA-TAC-GGC-GCC-TAC-GGA-TAT-CAA-TGG-CAT-3'  
\* \* \*
8. Mutants made: All 19 at 217

FIG.-22

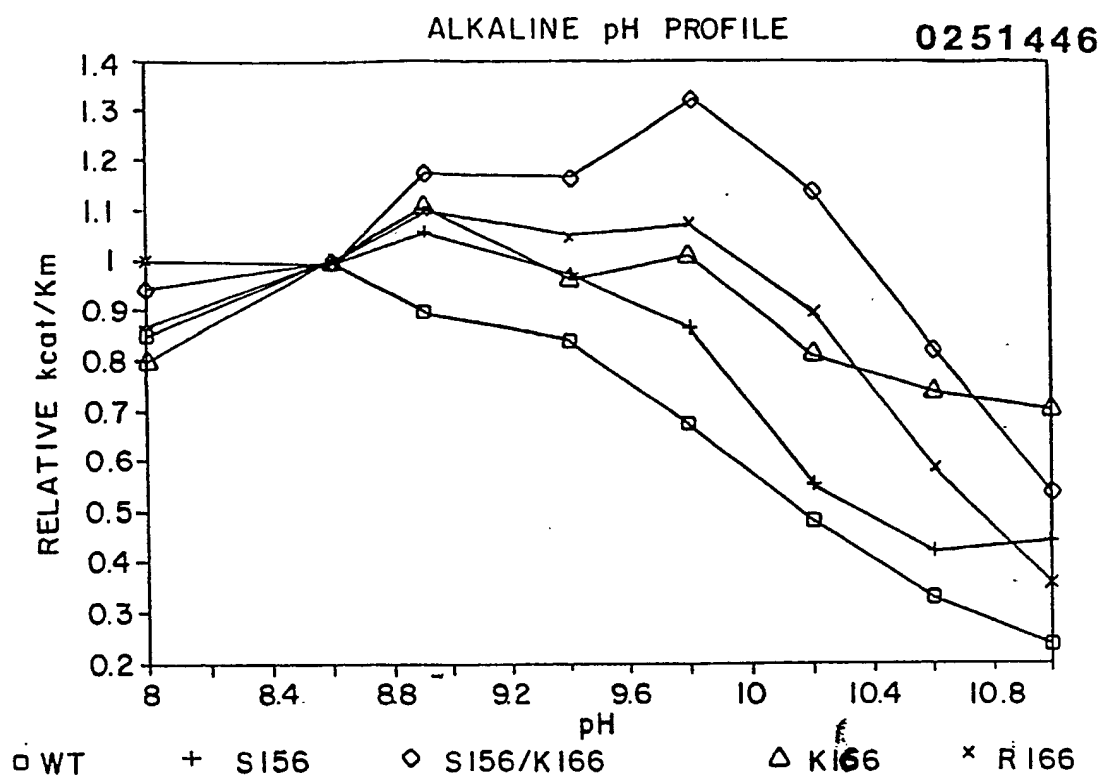


FIG. - 23A

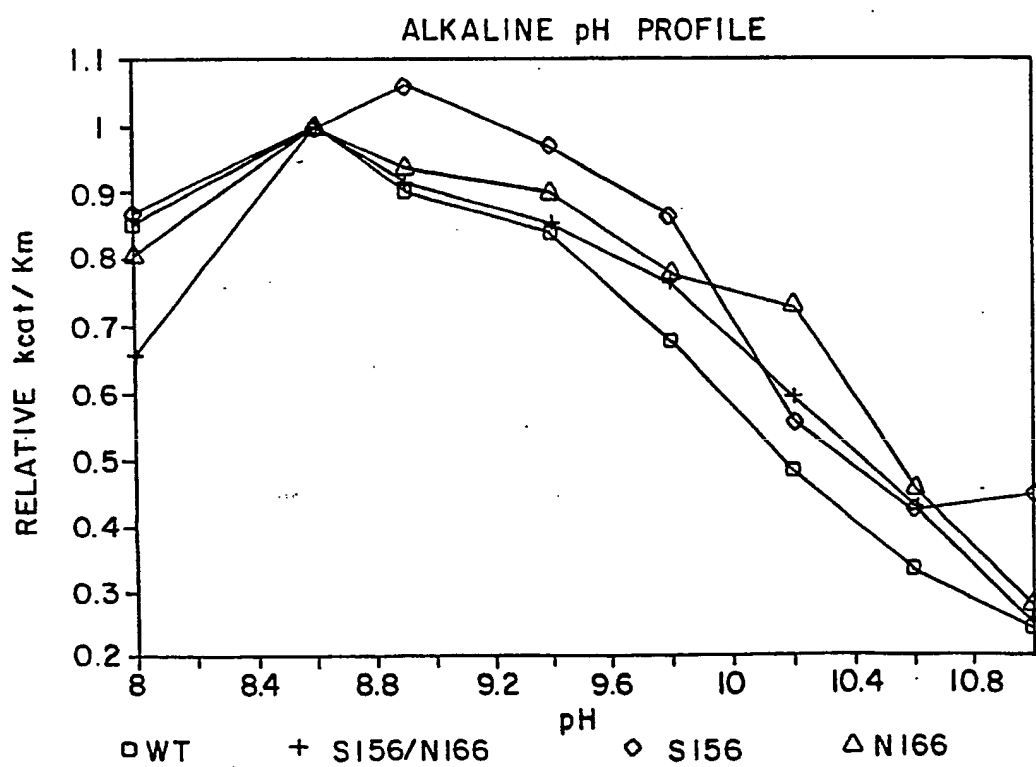


FIG. - 23B

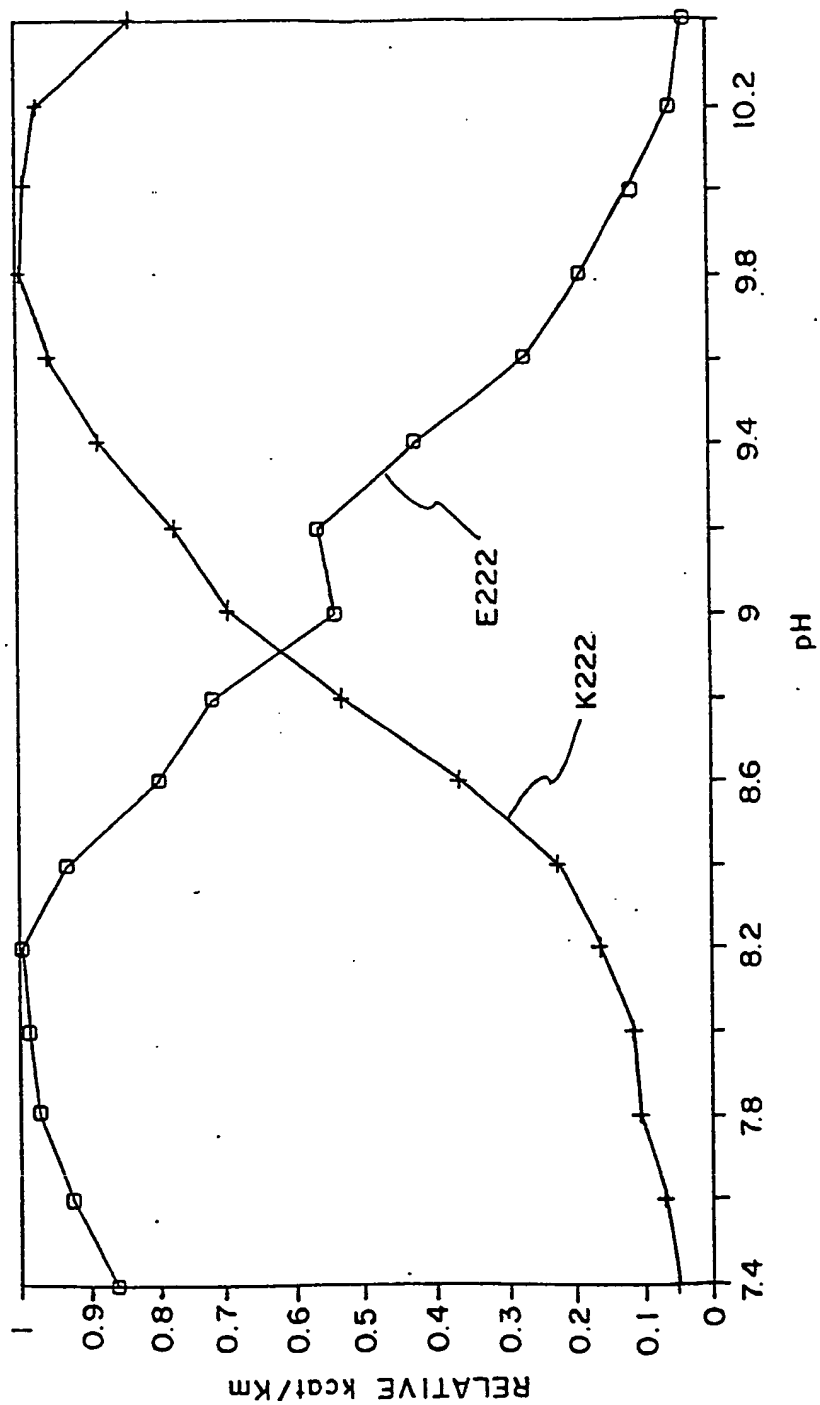
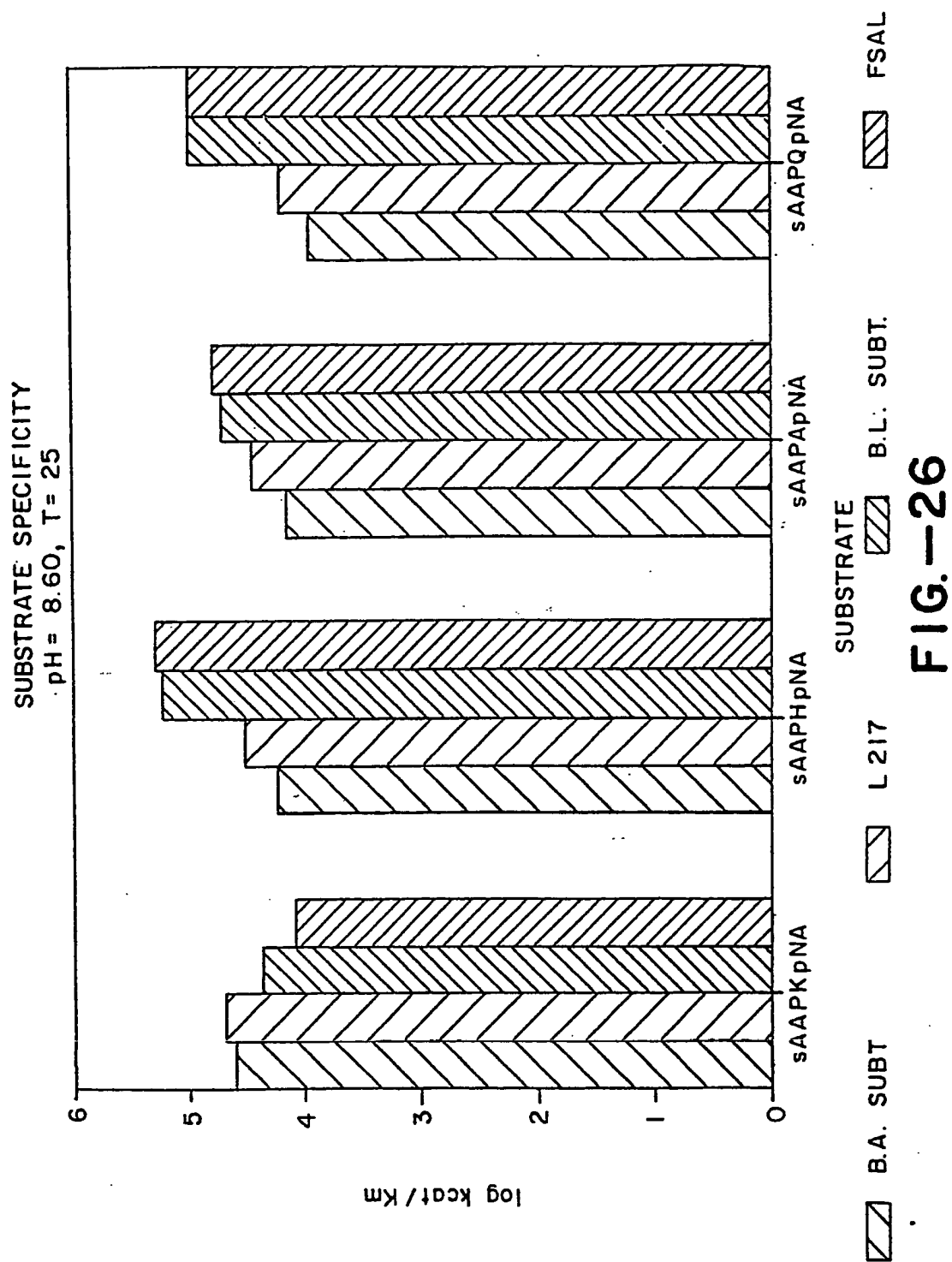


FIG.-24

1. Codon number: 91 95 100  
Tyr-Ala-Val-Lys-Val-Leu-Gly-Ala-Asp-Gly-Ser
2. Wild type amino acid sequence:  
5' -TAC-GCT-GTA-AAA-GTT-CTC-GGT-GGT-GAC-GGT-TCC  
ATG-CGA-CAT-TTT-CAA-GAG-CCA-CGA-CTG-CCA-AGG-5'
3. Wild type DNA sequence:  
5' -TAC-GCG-T-CTC-GGT-GCA-GAC-GGT-TCC  
ATG-CGC-A-GAG-CCA-CGT-CTG-CCA-AGG-5'  
Mu I Pst I
4. pΔ95:  
5' -TA \* PGAC-GGT-TCC  
ATG-CGCp A-CGT-CTG-CCA-AGG-5'
5. pΔ95 cut with *MuI* and *Pst* I:  
5' -TAC-GCG-GTA-AAA-GTT-CTC-GGT-GCA-GAC-GGT-TCC  
ATG-CGC-CAT-TTT-CAA-GAG-CCA-CGT-CTG-CCA-AGG-5'
6. Cut pΔ95 ligated with cassettes:  
5' -CA-TCA-CTT-TAC-GCG-T-CTC-GGT-GCA-GAC-GGT-TCC
7. Mutagenesis primer for pΔ95:  
5' -CA-TCA-CTT-TAC-GCG-T-CTC-GGT-GCA-GAC-GGT-TCC
8. Mutants made: C94, C95, D96

FIG.-25

0251446





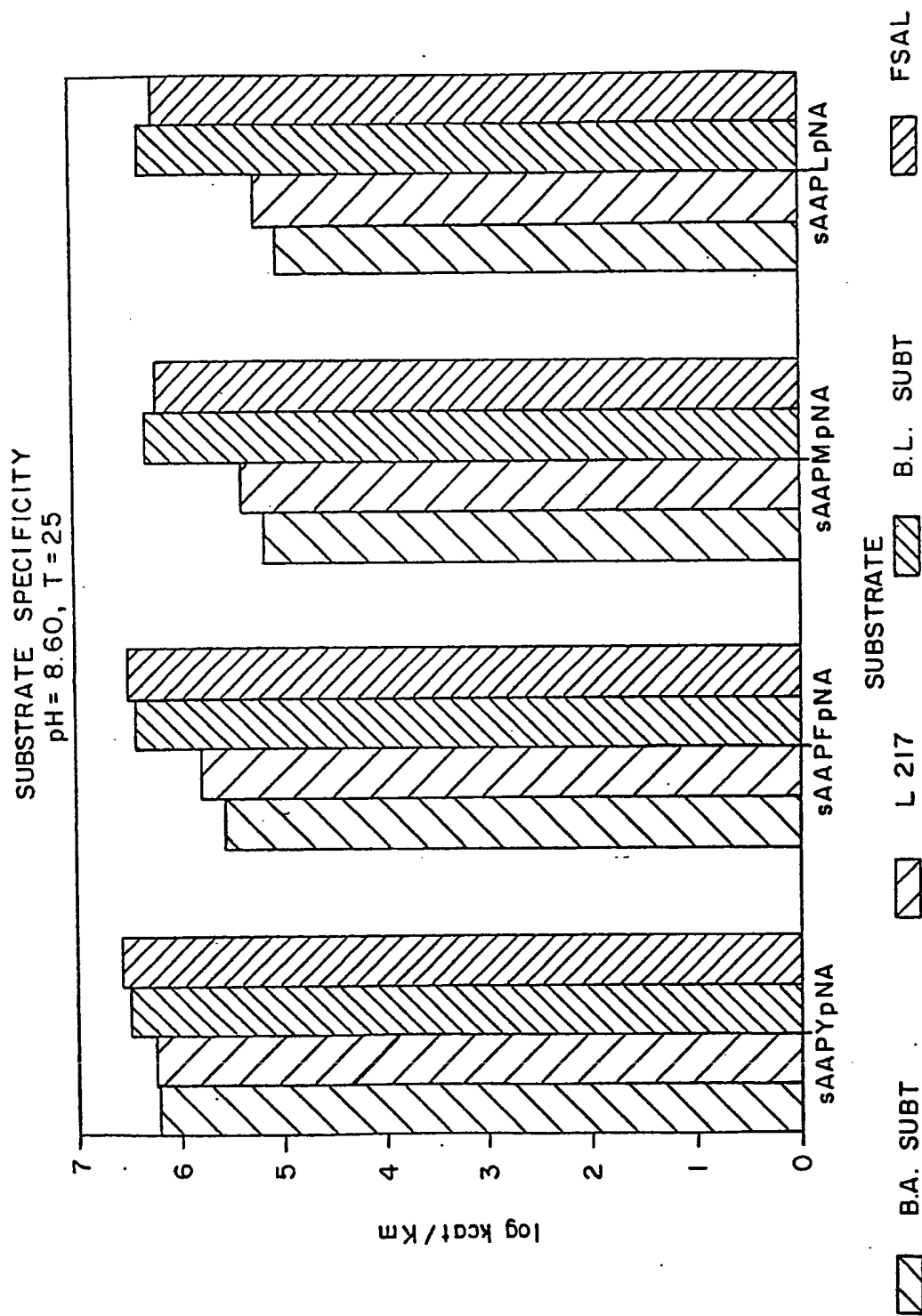


FIG.-27

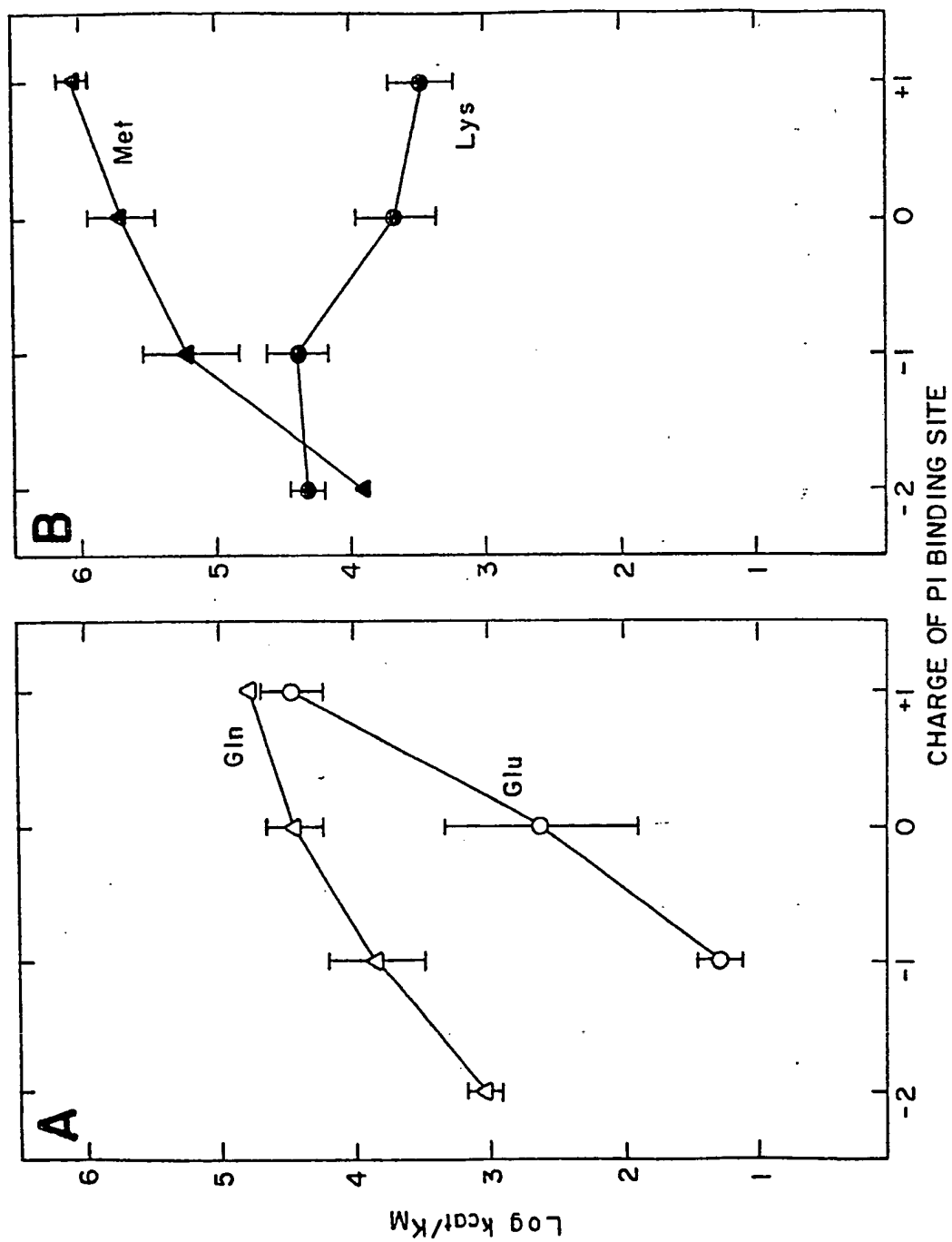


FIG.-28

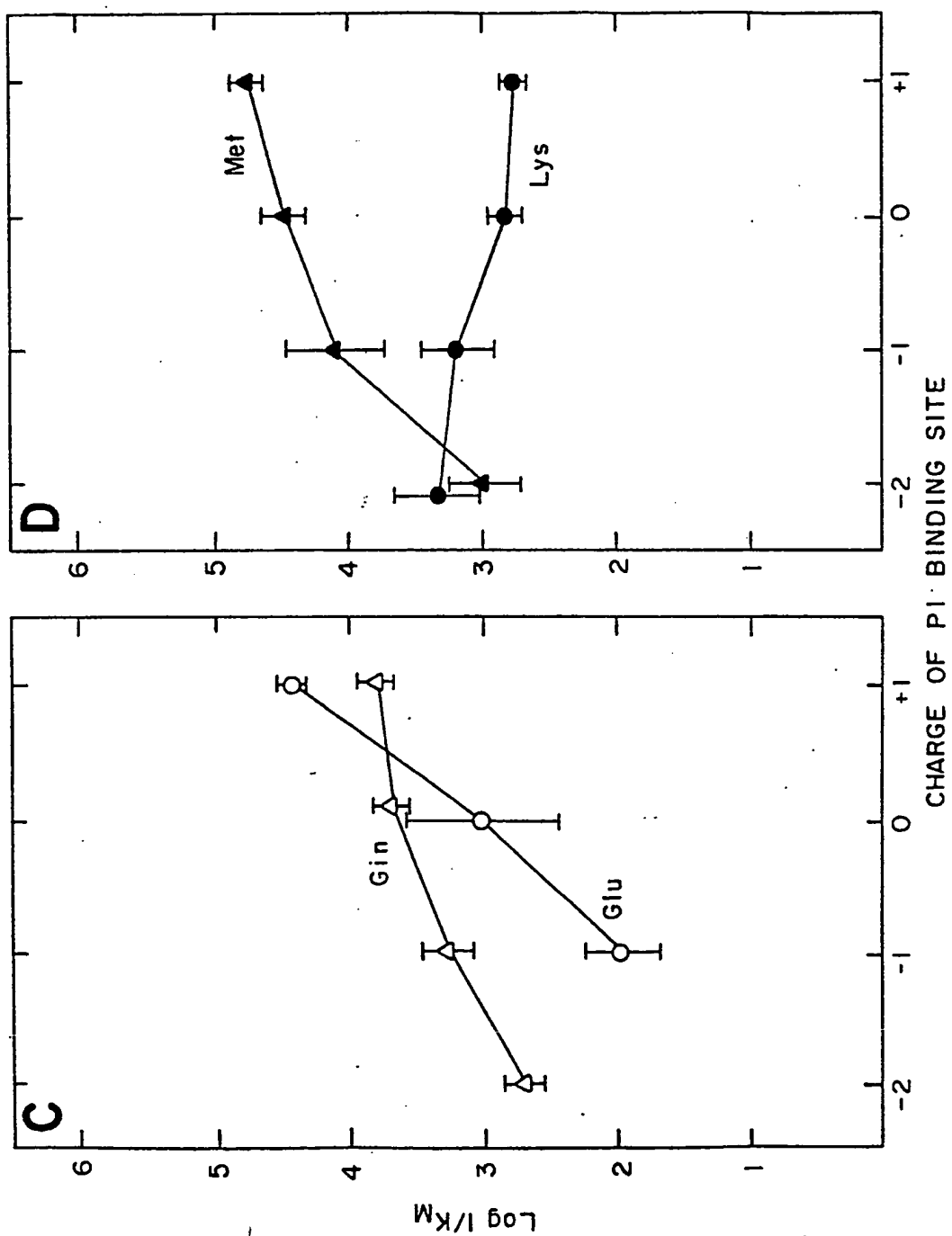


FIG.-28

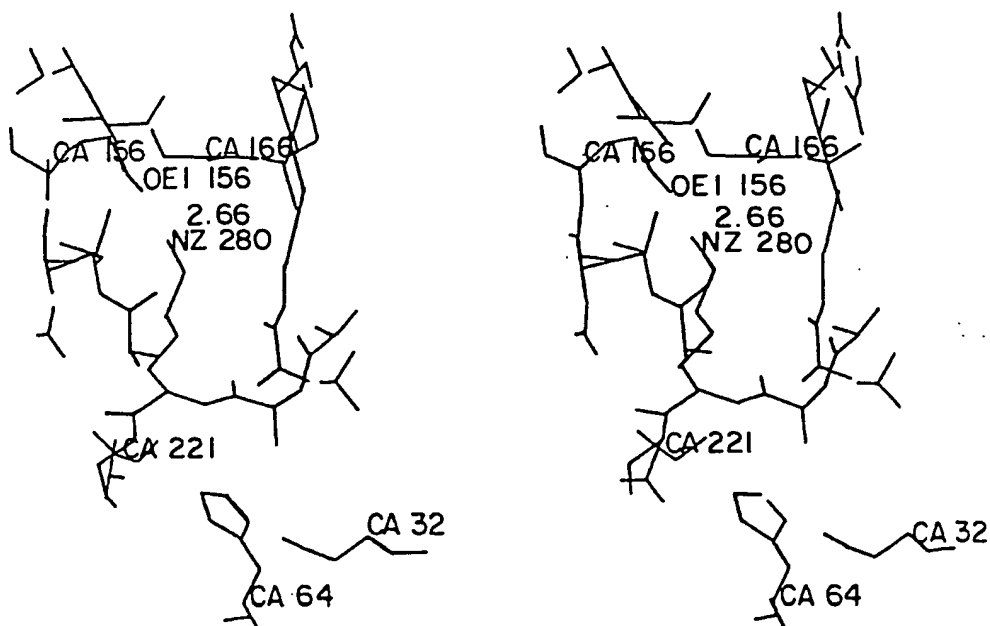


FIG.—29A

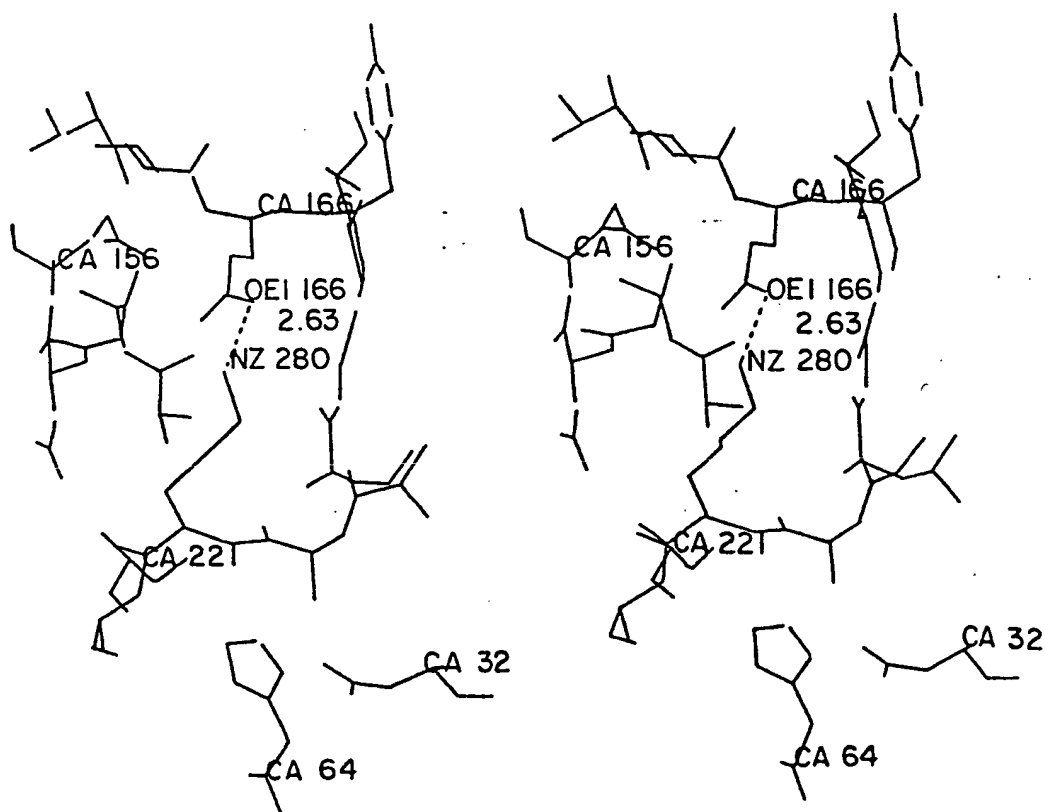


FIG.—29B

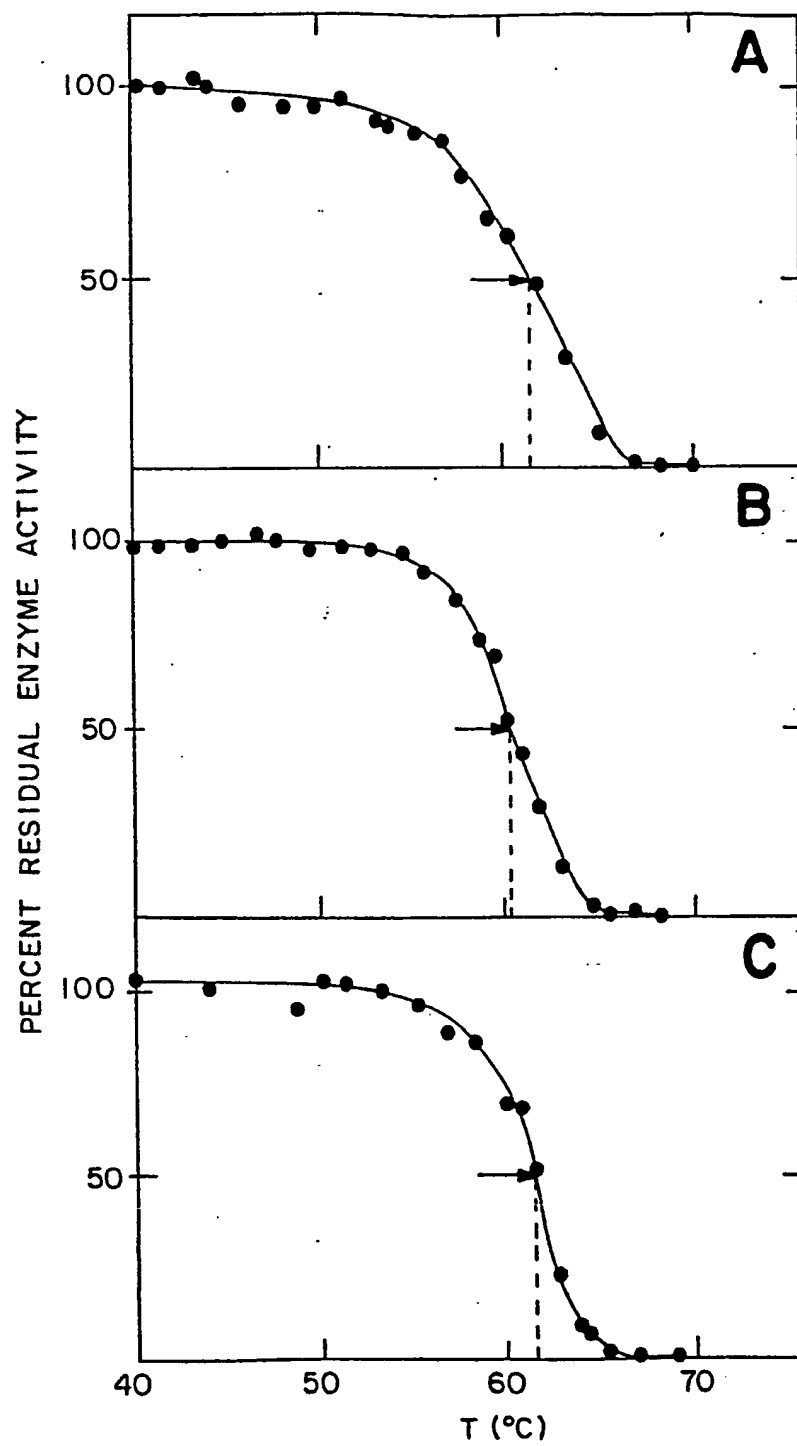


FIG.—30

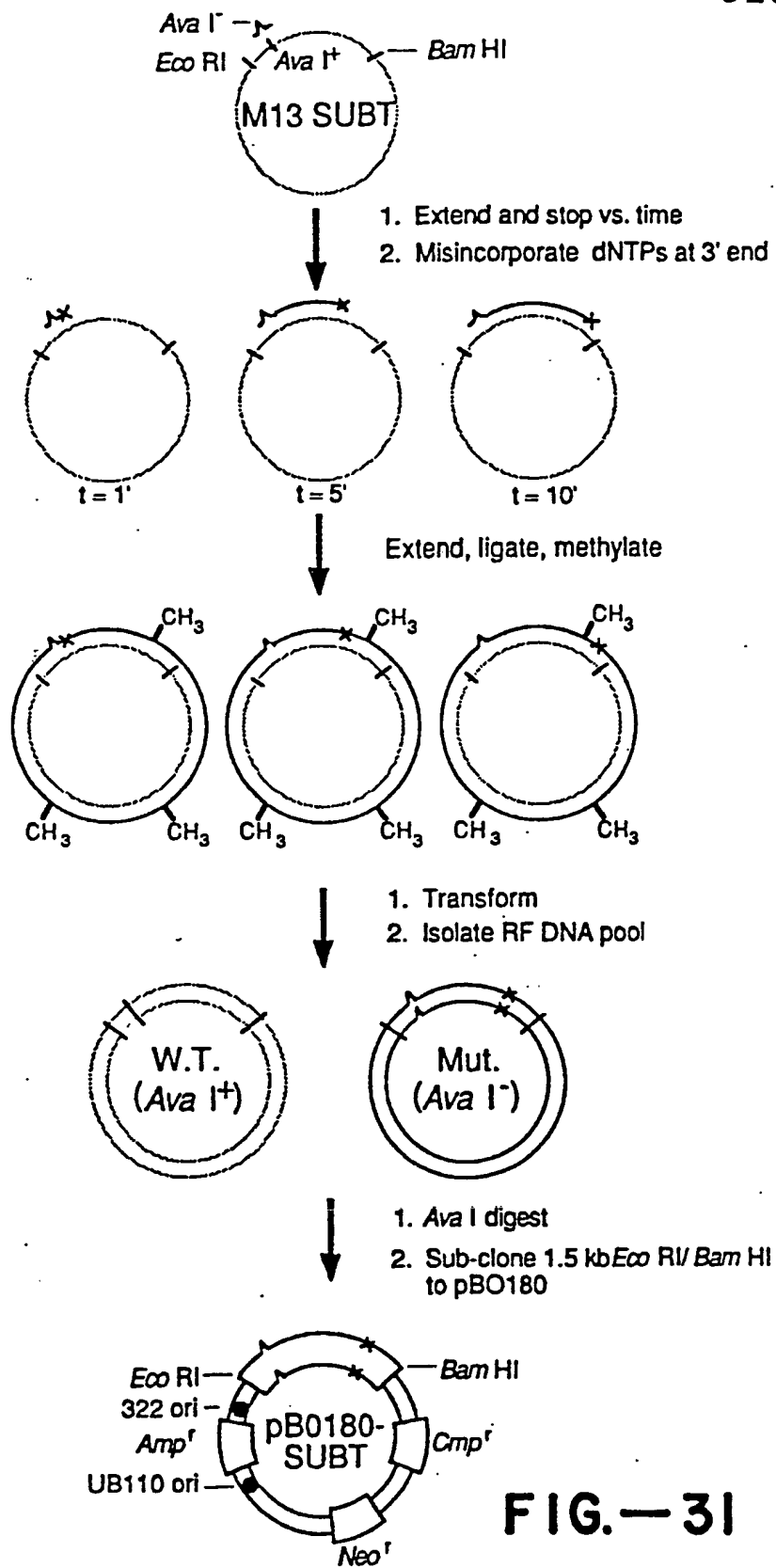


FIG.—31

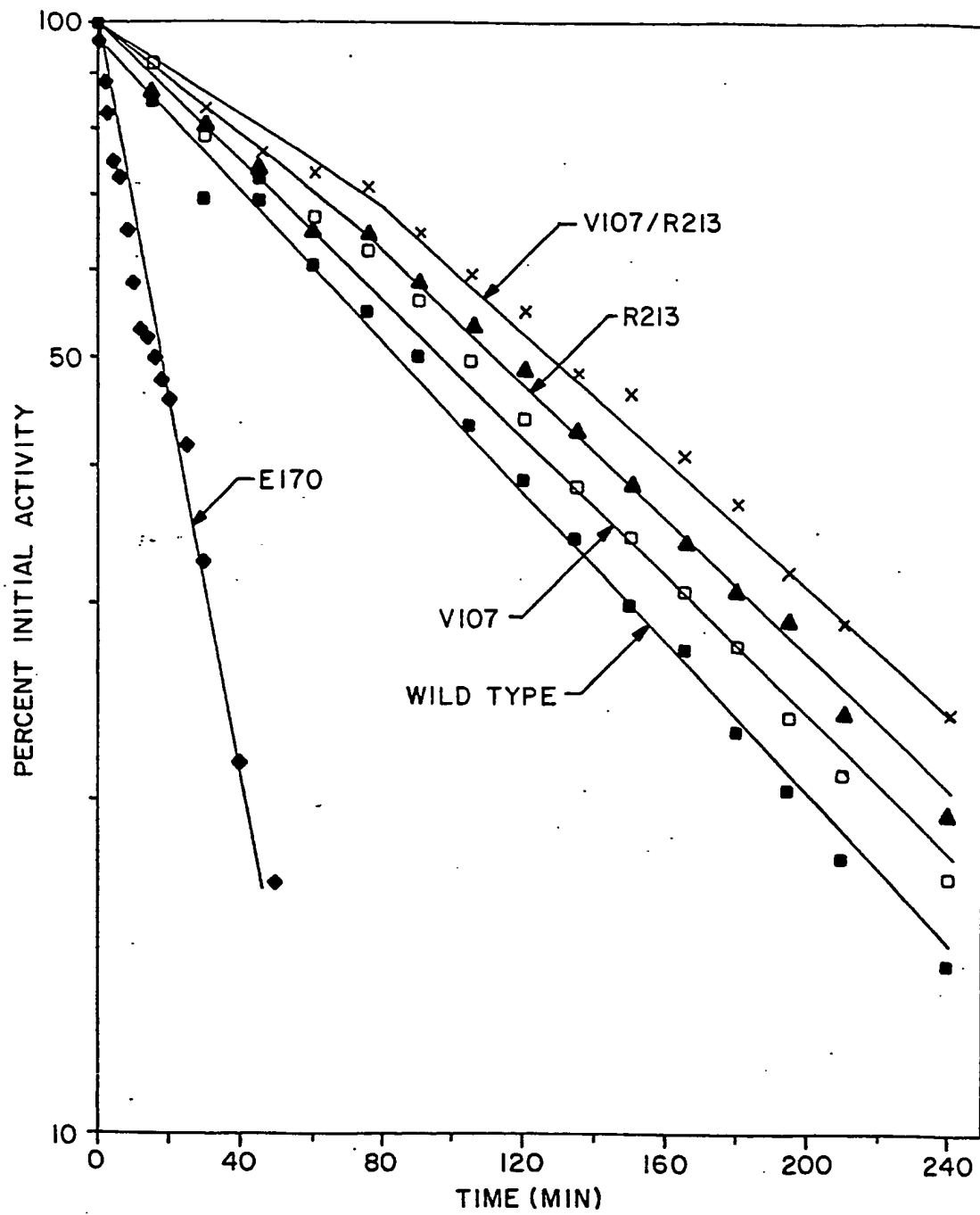


FIG.-32

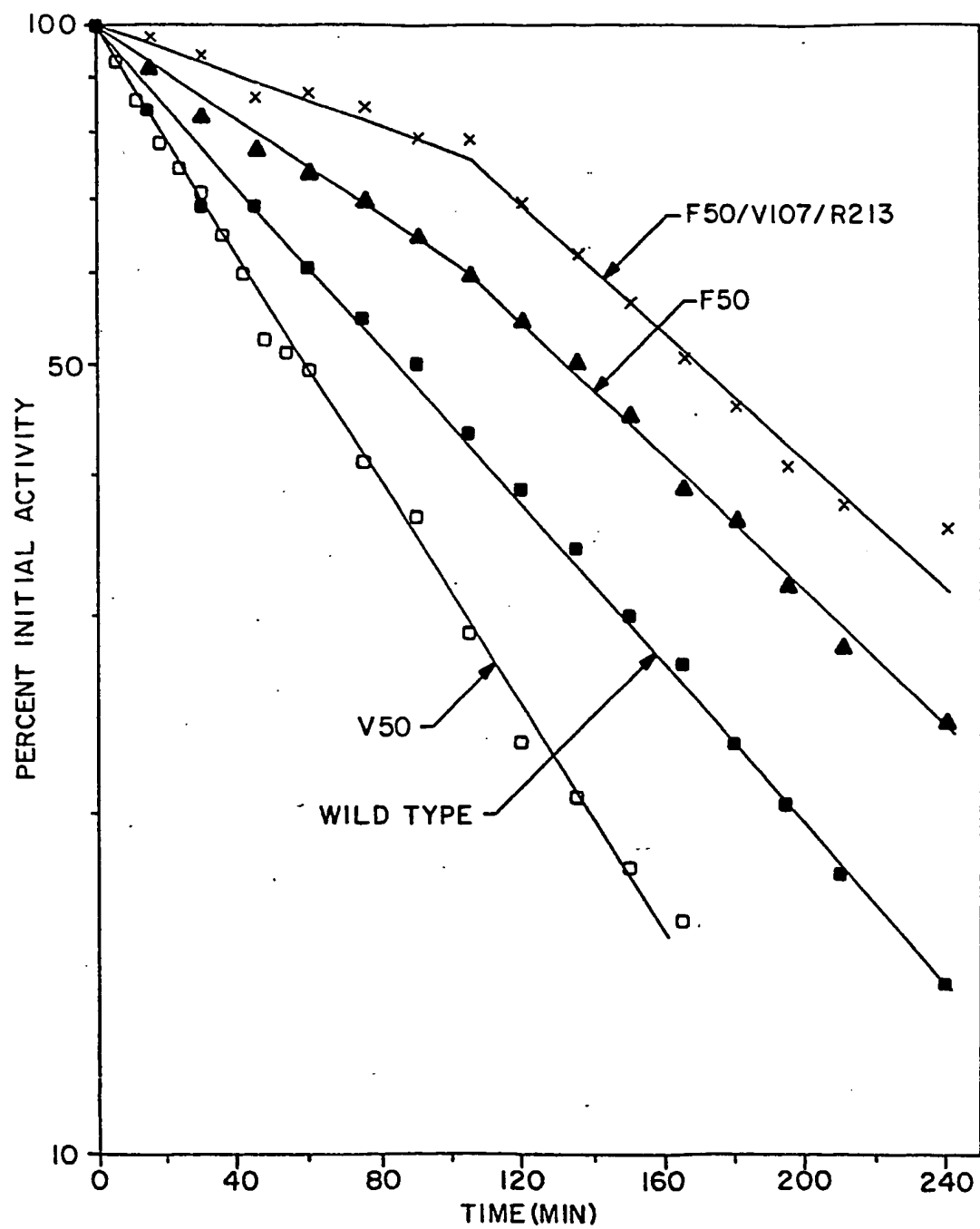


FIG.-33



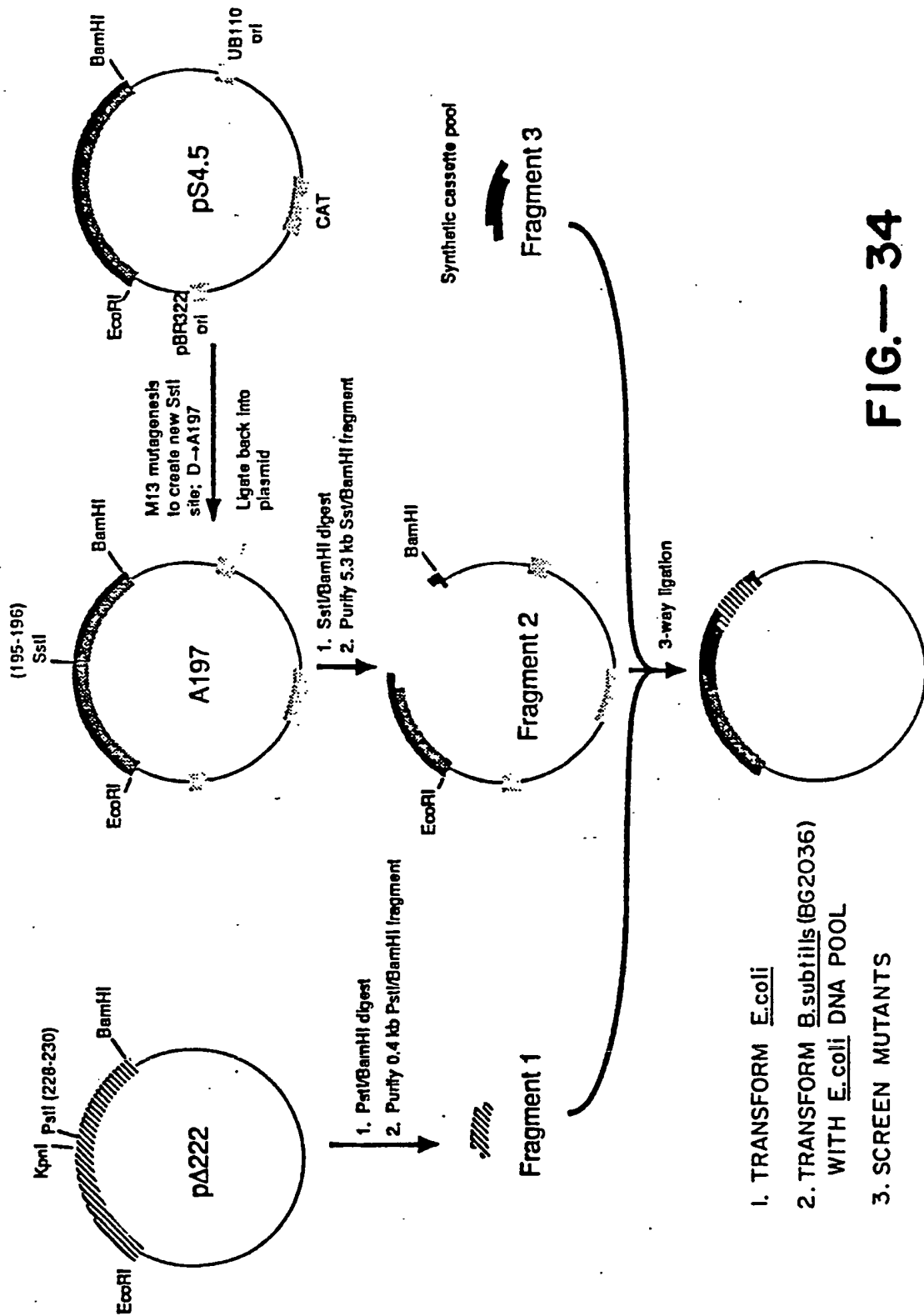


FIG.— 34

0251446

	195	200	206
W.T.A.A.:	Glu	Leu Asp Val Met Ala Pro Gly Val Ser Ile Gln	
W.T. DNA:	GAG CTT GAT GTC ATG GCA CCT GGC GTA TCT ATC CAA		
	CTC GAA CTA CAG TAC CGT GGA CCG CAT AGA TAG GTT		
pΔ222 DNA:	GAG CTT GAT GTC ATG GCA CCT GGC GTA TCT ATC CAA		
	CTC GAA CTA CAG TAC CGT GGA CCG CAT AGA TAG GTT		
A197 DNA:	<u>GAG CTC</u> <sup>*</sup> GCA <sup>**</sup> GTC ATG GCA CCT GGC GTA TCT ATC CAA		
	CTC GAG CGT CAG TAC CGT GGA CCG CAT AGA TAG GTT		
	<i>SstI</i>		
Fragments from pΔ222 and A197 cut w/ <i>PstI</i> , <i>SstI</i> :	GAG-CT		
	Cp		
	<sup>*</sup>		
pΔ222, A197 cut & ligated w/ oligodeoxy- nucleotide pools:	<u>GAG CTC</u> GAT GTC ATG GCA CCT GGC GTA TCT ATC CAA		
	CTC GAG CTA CAG TAC CGT GGA CCG CAT AGA TAG GTT		
	<i>SstI</i>		
	207	210	218
W.T.A.A.:	Ser Thr Leu Pro Gly Asn Lys Tyr Gly Ala Tyr Asn		
W.T. DNA:	AGC ACG CTT CCT GGA AAC AAA TAC GGG GCG TAC AAC		
	TCG TGC GAA GGA CCT TTG TTT ATG CCC CGC ATG TTG		
pΔ222 DNA:	AGC ACG CTT CCT GGA AAC AAA TAC GGG GCG TAC AAC		
	TCG TGC GAA GGA CCT TTG TTT ATG CCC CGC ATG TTG		
A197 DNA:	AGC ACG CTT CCT GGA AAC AAA TAC GGG GCG TAC AAC		
	TCG TGC GAA GGA CCT TTG TTT ATG CCC CGC ATG TTG		
	<sup>*</sup> <sup>*</sup>		
Fragments from pΔ222 and A197 cut w/ <i>PstI</i> , <i>SstI</i> :	AGC ACG CTT <u>CCC GGG</u> AAC AAA TAC GGG GCG TAC AAC		
	TCG TGC GAA <u>GGG CCC</u> TTG TTT ATG CCC CGC ATG TTG		
	<i>SmaI</i>		
	219	220	230
W.T.A.A.:	Gly Thr Ser Met Ala Ser Pro His Val Ala Gly Ala		
W.T. DNA:	GGT ACG TCA ATG GCA TCT CCG CAC GTT GCC GGA GCG-3'		
	CCA TGC AGT TAC CGT AGA GGC GTG CAA CGG CCT CGC-5'		
pΔ222 DNA:	<u>GGT ACC</u> <sup>*</sup> TCA-----CG CAC <u>GCT GCA</u> <sup>*</sup> GGA GCG-3'		
	CCA TGG AGT-----GC GTG CGA CGT CCT CGC-5'		
	<i>KpnI</i>	<i>PstI</i>	
A197 DNA:	GGT ACG TCA ATG GCA TCT CCG CAC GTT GCC GGA GCG-3'		
	CCA TGG AGT TAC CGT AGA GGC GTG CAA GTG CCT CGC-5'		
	<sup>*</sup>		
Fragments from pΔ222 and A197 cut w/ <i>PstI</i> , <i>SstI</i> :	<u>GGT ACC</u> TCA ATG GCA TCT CCG CAC GTT GCA GGA GCG-3'		
	CCA TGG AGT TAC CGT AGA GGC GTG CAA CGT CCT CGC-5'		
	<i>KpnI</i>	<i>PstI</i> destroyed	
	<sup>*</sup>		
pΔ222, A197 cut & ligated w/ oligodeoxy- nucleotide pools:	<u>GGT ACC</u> TCA ATG GCA TCT CCG CAC GTT GCA GGA GCG-3'		
	CCA TGG AGT TAC CGT AGA GGC GTG CAA CGT CCT CGC-5'		
	<i>KpnI</i>	<i>PstI</i> destroyed	

Oligodeoxynucleotide pools synthesized with 2% contaminating nucleotides in each cycle to give  
 -15% of pool with 0 mutations, -28% of pool with single mutations, and  
 -57% of pool with 2 or more mutations, according to the general formula  $f = \frac{\mu^n}{n!} e^{-\mu}$ .

FIG.—35

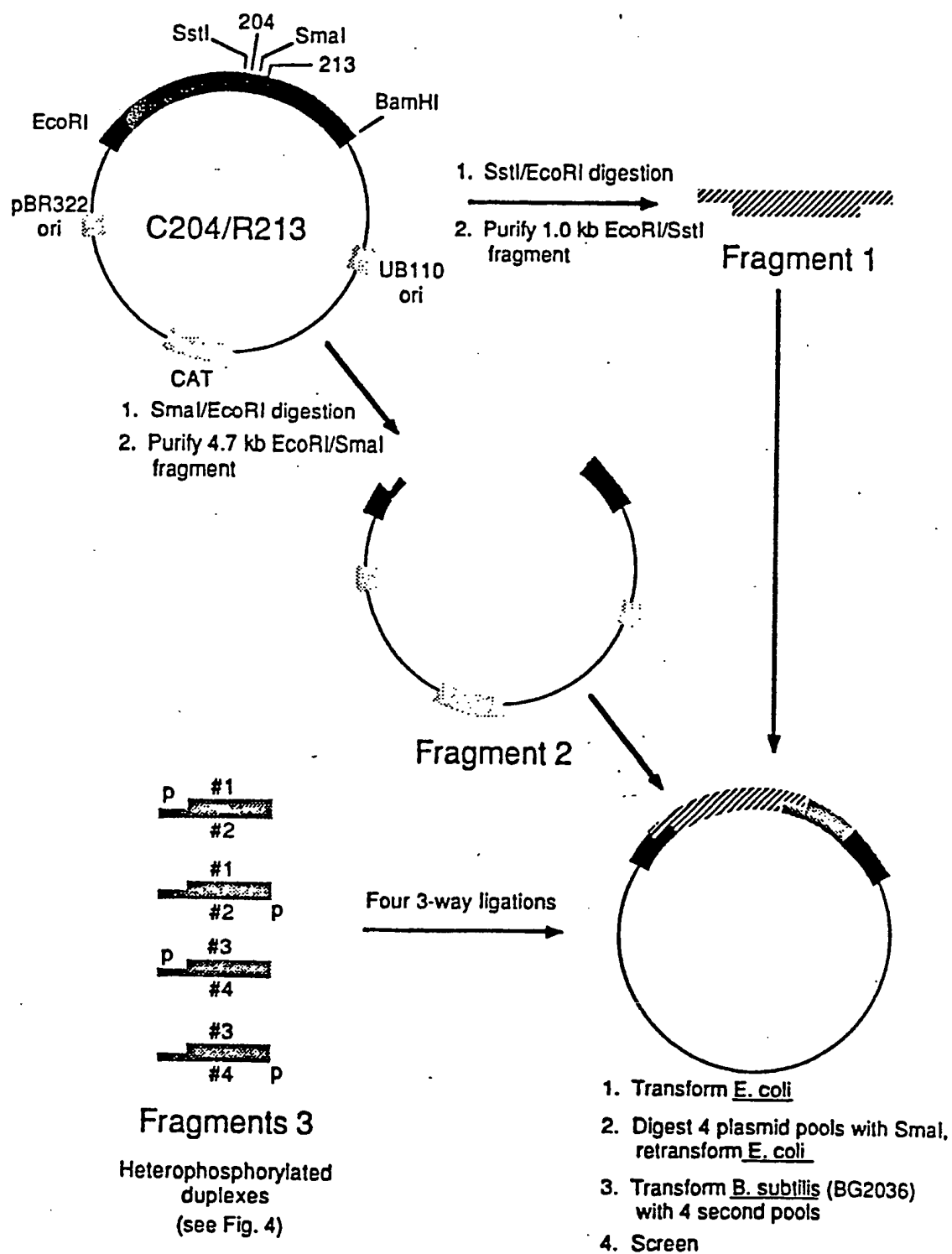


FIG.—36

Wild type A.A.: 195 Glu Leu Asp Val Met Ala Pro Gly Val Ser Ile Glu Ser Thr Leu Pro Gly Asn Lys 200 204 210 213

Wild type DNA: 5'-GAG CTT GAT GTC ATG GCA CCT GGC GTA TCT ATC CAA AGC ACG CTT CCT GGA AAC AAA-3' 3'-CTC GAA CTA CAG TAC CGT GGA CCG CAT AGA TAG GTT TCG TGC GAA GGA CCT TTG TTT-5'

C204/R213 DNA: 5'-GAG CTC GAT GTC ATG GCA CCT GGC GTA TGT ATC CAA AGC ACG CTT CCC GGG AAC AGA-3' 3'-CTC GAG CTA CAG TAC CGT GGA CCG CAT ACA TAG GTT TCG TGC GAA GGA CCC TTG TCT-5' SstI SmaI

C204/R213 cut with SstI and SmaI: 5'-GAG CT 3'-C GGG AAC AGA-3' CCC TTG TCT-5'

C204/R213 cut and ligated with oligodeoxynucleotide pools: 5'-GAG CTC GAT GTC ATG GCA CCT GGC GTA ATC CAG TCG ACG CTT CCT GGG AAC AGA-3' 3'-CTC GAG CTA CAG TAC CGT GGA CCG CAT TAG GTC AGC ACG TGC GAA GGA CCC TTG TCT-5' SstI SmaI

W, R, R, or G ← NGG or NCC → S, P, T or A

Stop, Y, H, Q, N, K, D or E ← [G]<sub>TN</sub> or [G]<sub>AN</sub> → L, F, I, V or M

#1 #2 #3 #4

0251446

FIG.—37

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

### **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☒ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**